

# (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  
2018 年 11 月 8 日 (08.11.2018)



(10) 国际公布号  
**WO 2018/201772 A1**

- (51) 国际专利分类号:  
**G06F 19/00** (2018.01)
- (21) 国际申请号: PCT/CN2018/076149
- (22) 国际申请日: 2018 年 2 月 10 日 (10.02.2018)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
201710313520.1 2017年5月5日 (05.05.2017) 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八卦岭工业区平安大厦六楼, Guangdong 518000 (CN)。
- (72) 发明人: 赵清源 (ZHAO, Qingyuan); 中国广东省深圳市福田区八卦岭工业区平安大厦六楼, Guangdong 518000 (CN)。 韦邕 (WEI, Yong); 中国广东省深圳市福田区八卦岭工业区平安大厦六楼, Guangdong 518000 (CN)。 吕梓燊 (LV, ZiShen); 中国广东省深圳市福田区八卦岭工业区

平安大厦六楼, Guangdong 518000 (CN)。 徐亮 (XU, Liang); 中国广东省深圳市福田区八卦岭工业区平安大厦六楼, Guangdong 518000 (CN)。 肖京 (XIAO, Jing); 中国广东省深圳市福田区八卦岭工业区平安大厦六楼, 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园岭街道八卦四路 10 号中浩大厦 1528-1530 室于志光, 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, 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,

(54) Title: METHOD AND SYSTEM FOR INFERRING POTENTIAL DISEASE FROM MEDICAL TEXT, AND READABLE STORAGE MEDIUM

(54) 发明名称: 医疗文本的潜在疾病推断方法、系统及可读存储介质

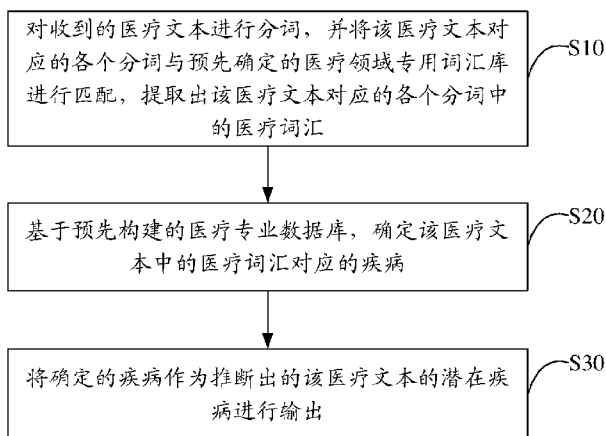


图 1

- S10 Perform segmentation on a received medical text, perform matching between respective words corresponding to the medical text and a pre-determined medical-specific terminology base, and extract a medical terminology from the respective words corresponding to the medical text
- S20 Determine, on the basis of a pre-established medical specialty database, a disease corresponding to the medical terminology in the medical text
- S30 Output the determined disease as a potential disease inferred from the medical text

(57) Abstract: Provided are a method and system for inferring a potential disease from a medical text, and a readable storage medium. The method comprises: performing segmentation on a received medical text, performing matching between respective words corresponding to the medical text and a pre-determined medical-specific terminology base, and extracting a medical terminology from the respective words corresponding to the medical text (S10); determining, on the basis of a pre-established medical specialty database, a disease corresponding to the medical terminology in the medical text (S20); and outputting the determined disease as a potential disease inferred from the medical text (S30). The method enables accurate and highly efficient inference of a potential disease from a medical text.

(57) 摘要: 一种医疗文本的潜在疾病推断方法、系统及可读存储介质, 该方法包括: 对收到的医疗文本进行分词, 并将该医疗文本对应的各个分词与预先确定的医疗领域专用词汇库进行匹配, 提取出该医疗文本对应的各个分词中的医疗词汇 (S10); 基于预先构建的医疗专业数据库, 确定该医疗文本中的医疗词汇对应的疾病 (S20); 将确定的疾病作为推断出的该医疗文本的潜在疾病进行输出 (S30)。该方法准确、高效地推断出医疗文本的潜在疾病。

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,  
US, UZ, VC, VN, 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))。

## 医疗文本的潜在疾病推断方法、系统及可读存储介质

### 优先权申明

本申请基于巴黎公约申明享有2017年5月5日递交的申请号为CN 2017103135201、名称为“医疗文本的潜在疾病推断方法、系统及可读存储介质”中国专利申请的优先权，该中国专利申请的整体内容以参考的方式结合在本申请中。

### 技术领域

本申请涉及计算机技术领域，尤其涉及一种医疗文本的潜在疾病推断方法、系统及可读存储介质。

### 背景技术

一般情况下，处理医疗文本的第一步都是推断潜在的疾病，进而才能进行接下来的诊断建议。现有技术中针对医疗文本的潜在疾病推断，只能根据医生的个人经验人工推断该医疗文本中的潜在疾病，效率较低，无法利用现有的医疗数据资源来进行潜在疾病的有效推断。

### 发明内容

本申请的主要目的在于提供一种医疗文本的潜在疾病推断方法、系统及可读存储介质，旨在准确、高效地推断出医疗文本的潜在疾病。

为实现上述目的，本申请第一方面提供一种医疗文本的潜在疾病推断方法，所述方法包括以下步骤：

A、对收到的医疗文本进行分词，并将该医疗文本对应的各个分词与预先确定的医疗领域专用词汇库进行匹配，提取出该医疗文本对应的各个分词中的医疗词汇；

B、基于预先构建的医疗专业数据库，确定该医疗文本中的医疗词汇对应的疾病；其中，所述医疗专业数据库中包含不同类型疾病与医疗词汇的映射关系；

C、将确定的疾病作为推断出的该医疗文本的潜在疾病进行输出。

此外，为实现上述目的，本申请第二方面还提供一种医疗文本的潜在疾病推断系统，所述医疗文本的潜在疾病推断系统包括：

分词提取模块，用于对收到的医疗文本进行分词，并将该医疗文本对应的各个分词与预先确定的医疗领域专用词汇库进行匹配，提取出该医疗文本对应的各个分词中的医疗词汇；

确定模块，用于基于预先构建的医疗专业数据库，确定该医疗文本中的医疗词汇对应的疾病；其中，所述医疗专业数据库中包含不同类型疾病与医疗词汇的映射关系；

输出模块,用于将确定的疾病作为推断出的该医疗文本的潜在疾病进行输出。

进一步地,为实现上述目的,本申请第三方面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存储有医疗文本的潜在疾病推断系统,所述医疗文本的潜在疾病推断系统可被至少一个处理器执行,以使所述至少一个处理器执行如上述的医疗文本的潜在疾病推断方法的步骤。

本申请提出的医疗文本的潜在疾病推断方法、系统及可读存储介质,通过对收到的医疗文本进行分词,提取出该医疗文本对应的各个分词中的医疗词汇;并基于预先构建的包含不同疾病与医疗词汇的映射关系的医疗专业数据库,确定该医疗文本中的医疗词汇对应的疾病,以作为推断出的该医疗文本的潜在疾病。由于能根据各种医疗数据资源构建不同疾病与医疗词汇的映射关系,并根据医疗文本中的医疗词汇找到与之映射的疾病,相比根据医生个人经验进行人工推断,效率更高且准确率更高。

#### **附图说明**

图1为本申请医疗文本的潜在疾病推断方法第一实施例的流程示意图;

图2为本申请医疗文本的潜在疾病推断方法第二实施例的流程示意图;

图3为本申请医疗文本的潜在疾病推断系统10较佳实施例的运行环境示意图;

图4为本申请医疗文本的潜在疾病推断系统第一实施例的功能模块示意图;

图5为本申请医疗文本的潜在疾病推断系统第二实施例的功能模块示意图。

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。

#### **具体实施方式**

为了使本申请所要解决的技术问题、技术方案及有益效果更加清楚、明白,以下结合附图和实施例,对本申请进行进一步详细说明。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。

本申请提供一种医疗文本的潜在疾病推断方法。

参照图 1，图 1 为本申请医疗文本的潜在疾病推断方法一实施例的流程示意图。

在一实施例中，该医疗文本的潜在疾病推断方法包括：

步骤 S10，对收到的医疗文本进行分词，并将该医疗文本对应的各个分词与预先确定的医疗领域专用词汇库进行匹配，提取出该医疗文本对应的各个分词中的医疗词汇。

接收待诊断的医疗文本，如可接收用户通过浏览器、客户端 APP 等发送的待诊断的医疗文本。本实施例中，在收到医疗文本后，首先对收到的医疗文本进行分词处理。例如，可根据标点符号将医疗文本切分成一条条完整的语句，再对各个切分的语句进行分词处理，如可利用字符串匹配的分词方法对各个切分的语句进行分词处理，如正向最大匹配法，把一个切分的语句中的字符串从左至右来分词；或者，反向最大匹配法，把一个切分的语句中的字符串从右至左来分词；或者，最短路径分词法，一个切分的语句中的字符串里面要求切出的词数是最少的；或者，双向最大匹配法，正反向同时进行分词匹配。还可利用词义分词法对各个切分的语句进行分词处理，词义分词法是一种机器语音判断的分词方法，利用句法信息和语义信息来处理歧义现象来分词。还可利用统计分词法对各个切分的语句进行分词处理，从当前用户的历史搜索记录或大众用户的历史搜索记录中，根据词组的统计，会统计有些两个相邻的字出现的频率较多，则可将这两个相邻的字作为词组来进行分词。

对医疗文本完成分词处理后，将该医疗文本对应的各个分词与预先确定的医疗领域专用词汇库进行匹配，预先确定的医疗领域专用词汇库中可包括通用医药词典中的医药词库、根据大量医学文本（例如互联网上的开源医疗数据）中抽取得到的各种不同疾病对应的简介信息、症状信息、并发症信息、治疗药品信息或治疗科室信息中的医疗词汇，等等。该医疗领域专用词汇库可以是固定不变的，也可以是根据互联网上最新的开源医疗数据定期更新医疗领域专用词汇库中的医疗词汇。提取出该医疗文本对应的各个分词中与预先确定的医疗领域专用词汇库相匹配的医疗词汇，即可获取到该医疗文本中与其潜在疾病相关性较大的信息即提取出的医疗词汇。

步骤 S20，基于预先构建的医疗专业数据库，确定该医疗文本中的医疗词汇对应的疾病；其中，所述医疗专业数据库中包含不同类型疾病与医疗词汇的映射关系。

提取出该医疗文本对应的各个分词中与其潜在疾病相关性较大的医疗词汇后，基于预先构建的医疗专业数据库，确定该医疗文本中的医疗词汇对应的疾病。所述医疗专业数据库中包含不同类型疾病与医疗词汇（如根据大量医学文本中抽取得到的症状、药品、检查、科

室等信息词汇)的映射关系,如可根据网上开源数据和文本,构建医疗专业数据库,包含疾病及其对应的简介、症状、并发症、治疗药品、常见检查等专业信息。基于构建的不同疾病与医疗词汇的映射关系,可根据提取出的该医疗文本中的医疗词汇找到与之映射的疾病。

步骤 S30,将确定的疾病作为推断出的该医疗文本的潜在疾病进行输出。

根据提取出的该医疗文本中的医疗词汇确定出对应的疾病后,即可将确定的疾病作为推断出的该医疗文本的潜在疾病进行输出,以基于推断出的该医疗文本的潜在疾病来进行后续的诊断建议。经过实际应用中的医疗文本潜在疾病推断统计,通过本实施例中的潜在疾病推断方法得到的疾病标签准确率(人工审查没有明显错误)可以达到85%左右,能有效提高对医疗文本潜在疾病推断的准确率。

本实施例通过对收到的医疗文本进行分词,提取出该医疗文本对应的各个分词中的医疗词汇;并基于预先构建的包含不同疾病与医疗词汇的映射关系的医疗专业数据库,确定该医疗文本中的医疗词汇对应的疾病,以作为推断出的该医疗文本的潜在疾病。由于能根据各种医疗数据资源构建不同疾病与医疗词汇的映射关系,并根据医疗文本中的医疗词汇找到与之映射的疾病,相比根据医生个人经验进行人工推断,效率更高且准确率更高。

如图2所示,本申请第二实施例提出一种医疗文本的潜在疾病推断方法,在上述实施例的基础上,在上述步骤S10之前还包括:

步骤 S40,从预先确定的数据源获取医疗数据,从所述医疗数据中找出每一种疾病对应的一个或多个医疗词汇,并根据不同类型疾病与医疗词汇的映射关系建立医疗专业数据库。

本实施例中,在进行医疗文本的潜在疾病推断之前,先从预先确定的数据源获取医疗数据,以根据所述医疗数据中的不同类型疾病与医疗词汇的映射关系建立医疗专业数据库。该医疗数据可以从现有的医疗数据库中获取的各种疾病的权威解释,包括其对应的简介、症状、并发症、治疗药品、常见检查等专业信息,也可以是各种药品对应的医疗信息,如药品主治的疾病类型等信息,该医疗数据也可以是通过网络爬虫等工具实时或者定时从互联网上的开源医疗数据源(例如,各大论坛上关于不同疾病的问答、讨论等,或各种最新的医疗案例、医疗问答文本等)获取的特定类型的信息(例如,不同疾病对应的治疗方案、治疗药物、所属科室、临床表现等)。从获取的医疗数据中找出每一种疾病对应的一个或多个医疗词汇,即可根据不同疾病与一个或多个医疗词汇的映射关系建立医疗专业数据库,以供后续基

于建立的医疗专业数据库来进行潜在疾病的推断。

进一步地，在其他实施例中，所述医疗专业数据库中还包括疾病对应的各个医疗词汇的权重，上述步骤 S20 可以包括：

基于预先构建的医疗专业数据库，找出该医疗文本中各个医疗词汇对应的疾病，并计算各个疾病对应的医疗词汇的权重加和，选择对应的医疗词汇的权重加和最高的疾病作为确定出的该医疗文本对应的疾病。

本实施例中，考虑到一种疾病对应的医疗词汇可能为一个或多个，一个医疗词汇对应的疾病也可能有一种或多种，例如，同一个症状可能会映射得到多个疾病、同一种药品也会治疗多种疾病。因此，在构建的医疗专业数据库中，还将不同医疗词汇赋予不同的权重，以便在基于构建的医疗专业数据库找出的医疗文本中各个医疗词汇有多个时，可计算各个疾病对应的医疗词汇的权重加和，选择对应的医疗词汇的权重加和最高的疾病作为确定出的该医疗文本对应的疾病。例如，可将某个疾病映射得到的权重加和作为推断该疾病的自信程度，选择自信程度最高的疾病作为最终结果，从而进一步提高对医疗文本潜在疾病推断的准确率。

进一步地，在其他实施例中，上述步骤 S10 中对收到的医疗文本进行分词处理的步骤包括：

根据正向最大匹配法将医疗文本中待处理的字符串与预先确定的医疗领域专用词汇库（例如，该医疗领域专用词汇库可以是通用医疗专业词库，也可以是可扩容的学习型医疗词库）进行匹配，得到第一匹配结果；

根据逆向最大匹配法将医疗文本中待处理的字符串与预先确定的医疗领域专用词汇库（例如，该医疗领域专用词汇库可以是通用医疗专业词库，也可以是可扩容的学习型医疗词库）进行匹配，得到第二匹配结果。其中，所述第一匹配结果中包含有第一数量的第一词组，所述第二匹配结果中包含有第二数量的第二词组；所述第一匹配结果中包含有第三数量的单字，所述第二匹配结果中包含有第四数量的单字。

若所述第一数量与所述第二数量相等，且所述第三数量小于或者等于所述第四数量，则输出该医疗文本对应的所述第一匹配结果（包括词组和单字）；

若所述第一数量与所述第二数量相等，且所述第三数量大于所述第四数量，则输出该医疗文本对应的所述第二匹配结果（包括词组和单字）；

若所述第一数量与所述第二数量不相等，且所述第一数量大于所述第二数量，则输出该医疗文本对应的所述第二匹配结果（包括词组和单字）；

若所述第一数量与所述第二数量不相等，且所述第一数量小于所述第二数量，则输出该医疗文本对应的所述第一匹配结果（包括词组和单字）。

本实施例中采用双向匹配法来对医疗文本进行分词处理，通过正反向同时进行分词匹配来分析医疗文本待处理的字符串中前后组合内容的粘性，由于通常情况下词组能代表核心观点信息的概率更大，即词组更有可能是该医疗文本中的医疗词汇。因此，通过正反向同时进行分词匹配找出单字数量更少，词组数量更多的分词匹配结果，以作为医疗文本的分词结果，从而提高分词的准确性，以更加准确地提取出该医疗文本中的医疗词汇。

本申请进一步提供一种医疗文本的潜在疾病推断系统。请参阅图3，是本申请医疗文本的潜在疾病推断系统10较佳实施例的运行环境示意图。

在本实施例中，所述的医疗文本的潜在疾病推断系统10安装并运行于电子装置1中。该电子装置1可包括，但不限于，存储器11、处理器12及显示器13。图3仅示出了具有组件11-13的电子装置1，但是应理解的是，并不要求实施所有示出的组件，可以替代的实施更多或者更少的组件。

所述存储器11在一些实施例中可以是所述电子装置1的内部存储单元，例如该电子装置1的硬盘或内存。所述存储器11在另一些实施例中也可以是所述电子装置1的外部存储设备，例如所述电子装置1上配备的插接式硬盘，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还可以既包括所述电子装置1的内部存储单元也包括外部存储设备。所述存储器11用于存储安装于所述电子装置1的应用软件及各类数据，例如所述医疗文本的潜在疾病推断系统10的程序代码等。所述存储器11还可以用于暂时地存储已经输出或者将要输出的数据。

所述处理器12在一些实施例中可以是一中央处理器（Central Processing Unit, CPU），微处理器或其他数据处理芯片，用于运行所述存储器11中存储的程序代码或处理数据，例如执行所述医疗文本的潜在疾病推断系统10等。

所述显示器13在一些实施例中可以是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。所述显示器 13 用于显示在所述电子装置 1 中处理的信息以及用于显示可视化的用户界面,例如显示提取出的医疗文本中的医疗词汇、推断出的该医疗文本的潜在疾病等。所述电子装置 1 的部件 11-13 通过系统总线相互通信。

请参阅图 4,是本申请医疗文本的潜在疾病推断系统 10 第一实施例的功能模块图。在本实施例中,所述的医疗文本的潜在疾病推断系统 10 可以被分割成一个或多个模块,所述一个或者多个模块被存储于所述存储器 11 中,并由一个或多个处理器(本实施例为所述处理器 12)所执行,以完成本申请。例如,在图 4 中,所述的医疗文本的潜在疾病推断系统 10 可以被分割成分词提取模块 01、确定模块 02 及输出模块 03。本申请所称的模块是指能够完成特定功能的一系列计算机程序指令段,比程序更适合于描述所述语音识别系统 10 在所述电子装置 1 中的执行过程。以下描述将具体介绍所述分词提取模块 01、确定模块 02 及输出模块 03 的功能。

分词提取模块 01,用于对收到的医疗文本进行分词,并将该医疗文本对应的各个分词与预先确定的医疗领域专用词汇库进行匹配,提取出该医疗文本对应的各个分词中的医疗词汇;

接收待诊断的医疗文本,如可接收用户通过浏览器、客户端 APP 等发送的待诊断的医疗文本。本实施例中,在收到医疗文本后,首先对收到的医疗文本进行分词处理。例如,可根据标点符号将医疗文本切分成一条条完整的语句,再对各个切分的语句进行分词处理,如可利用字符串匹配的分词方法对各个切分的语句进行分词处理,如正向最大匹配法,把一个切分的语句中的字符串从左至右来分词;或者,反向最大匹配法,把一个切分的语句中的字符串从右至左来分词;或者,最短路径分词法,一个切分的语句中的字符串里面要求切出的词数是最少的;或者,双向最大匹配法,正反向同时进行分词匹配。还可利用词义分词法对各个切分的语句进行分词处理,词义分词法是一种机器语音判断的分词方法,利用句法信息和语义信息来处理歧义现象来分词。还可利用统计分词法对各个切分的语句进行分词处理,从当前用户的历史搜索记录或大众用户的历史搜索记录中,根据词组的统计,会统计有些两个相邻的字出现的频率较多,则可将这两个相邻的字作为词组来进行分词。

对医疗文本完成分词处理后,将该医疗文本对应的各个分词与预先确定的医疗领域专用词汇库进行匹配,预先确定的医疗领域专用词汇库中可包括通用医药词典中的医药词库、根据大量医学文本(例如互联网上的开源医疗数据)中抽取得到的各种不同疾病对应的简介信息、症状信息、并发症信息、治疗药品信息或治疗科室信息中的医疗词汇,等等。该医疗领域专用词汇库可以是固定不变的,也可以是根

据互联网上最新的开源医疗数据定期更新医疗领域专用词汇库中的医疗词汇。提取出该医疗文本对应的各个分词中与预先确定的医疗领域专用词汇库相匹配的医疗词汇，即可获取到该医疗文本中与其潜在疾病相关性较大的信息即提取出的医疗词汇。

确定模块 02，用于基于预先构建的医疗专业数据库，确定该医疗文本中的医疗词汇对应的疾病；其中，所述医疗专业数据库中包含不同类型疾病与医疗词汇的映射关系；

提取出该医疗文本对应的各个分词中与其潜在疾病相关性较大的医疗词汇后，基于预先构建的医疗专业数据库，确定该医疗文本中的医疗词汇对应的疾病。所述医疗专业数据库中包含不同类型疾病与医疗词汇（如根据大量医学文本中抽取得到的症状、药品、检查、科室等信息词汇）的映射关系，如可根据网上开源数据和文本，构建医疗专业数据库，包含疾病及其对应的简介、症状、并发症、治疗药品、常见检查等专业信息。基于构建的不同疾病与医疗词汇的映射关系，可根据提取出的该医疗文本中的医疗词汇找到与之映射的疾病。

输出模块 03，用于将确定的疾病作为推断出的该医疗文本的潜在疾病进行输出。

根据提取出的该医疗文本中的医疗词汇确定出对应的疾病后，即可将确定的疾病作为推断出的该医疗文本的潜在疾病进行输出，以基于推断出的该医疗文本的潜在疾病来进行后续的诊断建议。经过实际应用中的医疗文本潜在疾病推断统计，通过本实施例中的潜在疾病推断方法得到的疾病标签准确率（人工审查没有明显错误）可以达到85%左右，能有效提高对医疗文本潜在疾病推断的准确率。

本实施例通过对收到的医疗文本进行分词，提取出该医疗文本对应的各个分词中的医疗词汇；并基于预先构建的包含不同疾病与医疗词汇的映射关系的医疗专业数据库，确定该医疗文本中的医疗词汇对应的疾病，以作为推断出的该医疗文本的潜在疾病。由于能根据各种医疗数据资源构建不同疾病与医疗词汇的映射关系，并根据医疗文本中的医疗词汇找到与之映射的疾病，相比根据医生个人经验进行人工推断，效率更高且准确率更高。

如图5所示，本申请第二实施例提出一种医疗文本的潜在疾病推断系统，在上述实施例的基础上，还包括：

建立模块 04，用于从预先确定的数据源获取医疗数据，从所述医疗数据中找出每一种疾病对应的一个或多个医疗词汇，并根据不同类型疾病与医疗词汇的映射关系建立医疗专业数据库。

本实施例中，在进行医疗文本的潜在疾病推断之前，先从预先确

定的数据源获取医疗数据，以根据所述医疗数据中的不同类型疾病与医疗词汇的映射关系建立医疗专业数据库。该医疗数据可以从现有的医疗数据库中获取的各种疾病的权威解释，包括其对应的简介、症状、并发症、治疗药品、常见检查等专业信息，也可以是各种药品对应的医疗信息，如药品主治的疾病类型等信息，该医疗数据也可以是通过网络爬虫等工具实时或者定时从互联网上的开源医疗数据源（例如，各大论坛上关于不同疾病的问答、讨论等，或各种最新的医疗案例、医疗问答文本等）获取的特定类型的信息（例如，不同疾病对应的治疗方案、治疗药物、所属科室、临床表现等）。从获取的医疗数据中找出每一种疾病对应的一个或多个医疗词汇，即可根据不同疾病与一个或多个医疗词汇的映射关系建立医疗专业数据库，以供后续基于建立的医疗专业数据库来进行潜在疾病的推断。

进一步地，在其他实施例中，所述医疗专业数据库中还包括疾病对应的各个医疗词汇的权重，上述确定模块 02 还可以用于：

基于预先构建的医疗专业数据库，找出该医疗文本中各个医疗词汇对应的疾病，并计算各个疾病对应的医疗词汇的权重加和，选择对应的医疗词汇的权重加和最高的疾病作为确定出的该医疗文本对应的疾病。

本实施例中，考虑到一种疾病对应的医疗词汇可能为一个或多个，一个医疗词汇对应的疾病也可能有一种或多种，例如，同一个症状可能会映射得到多个疾病、同一种药品也会治疗多种疾病。因此，在构建的医疗专业数据库中，还将不同医疗词汇赋予不同的权重，以便在基于构建的医疗专业数据库找出的医疗文本中各个医疗词汇有多个时，可计算各个疾病对应的医疗词汇的权重加和，选择对应的医疗词汇的权重加和最高的疾病作为确定出的该医疗文本对应的疾病。例如，可将某个疾病映射得到的权重加和作为推断该疾病的自信程度，选择自信程度最高的疾病作为最终结果，从而进一步提高对医疗文本潜在疾病推断的准确率。

进一步地，在其他实施例中，上述分词提取模块 01 还用于：

根据正向最大匹配法将医疗文本中待处理的字符串与预先确定的医疗领域专用词汇库（例如，该医疗领域专用词汇库可以是通用医疗专业词库，也可以是可扩容的学习型医疗词库）进行匹配，得到第一匹配结果；

根据逆向最大匹配法将医疗文本中待处理的字符串与预先确定的医疗领域专用词汇库（例如，该医疗领域专用词汇库可以是通用医疗专业词库，也可以是可扩容的学习型医疗词库）进行匹配，得到第

二匹配结果。其中，所述第一匹配结果中包含有第一数量的第一词组，所述第二匹配结果中包含有第二数量的第二词组；所述第一匹配结果中包含有第三数量的单字，所述第二匹配结果中包含有第四数量的单字。

若所述第一数量与所述第二数量相等，且所述第三数量小于或者等于所述第四数量，则输出该医疗文本对应的所述第一匹配结果（包括词组和单字）；

若所述第一数量与所述第二数量相等，且所述第三数量大于所述第四数量，则输出该医疗文本对应的所述第二匹配结果（包括词组和单字）；

若所述第一数量与所述第二数量不相等，且所述第一数量大于所述第二数量，则输出该医疗文本对应的所述第二匹配结果（包括词组和单字）；

若所述第一数量与所述第二数量不相等，且所述第一数量小于所述第二数量，则输出该医疗文本对应的所述第一匹配结果（包括词组和单字）。

本实施例中采用双向匹配法来对医疗文本进行分词处理，通过正反向同时进行分词匹配来分析医疗文本待处理的字符串中前后组合内容的粘性，由于通常情况下词组能代表核心观点信息的概率更大，即词组更有可能是该医疗文本中的医疗词汇。因此，通过正反向同时进行分词匹配找出单字数量更少，词组数量更多的分词匹配结果，以作为医疗文本的分词结果，从而提高分词的准确性，以更加准确地提取出该医疗文本中的医疗词汇。

此外，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存储有医疗文本的潜在疾病推断系统，所述医疗文本的潜在疾病推断系统可被至少一个处理器执行，以使所述至少一个处理器执行如上述实施例中的医疗文本的潜在疾病推断方法的步骤，该医疗文本的潜在疾病推断方法的步骤 S10、S20、S30 等具体实施过程如上文所述，在此不再赘述。

需要说明的是，在本文中，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过程、方法、物品或者装置不仅包括那些要素，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过程、方法、物品或者装置所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括该要素的过程、方法、物品或者装置中还存在另外的相同要素。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件来实现，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质（如 ROM/RAM、磁碟、光盘）中，包括若干指令用以使得一台终端设备（可以是手机，计算机，服务器，空调器，或者网络设备等等）执行本申请各个实施例所述的方法。

以上参照附图说明了本申请的优选实施例，并非因此局限本申请的权利范围。上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。另外，虽然在流程图中示出了逻辑顺序，但是在某些情况下，可以以不同于此处的顺序执行所示出或描述的步骤。

本领域技术人员不脱离本申请的范围和实质，可以有多种变型方案实现本申请，比如作为一个实施例的特征可用于另一实施例而得到又一实施例。凡在运用本申请的技术构思之内所作的任何修改、等同替换和改进，均应在本申请的权利范围之内。

## 权利要求书

1、一种医疗文本的潜在疾病推断方法，其特征在于，所述方法包括以下步骤：

A、对收到的医疗文本进行分词，并将该医疗文本对应的各个分词与预先确定的医疗领域专用词汇库进行匹配，提取出该医疗文本对应的各个分词中的医疗词汇；

B、基于预先构建的医疗专业数据库，确定该医疗文本中的医疗词汇对应的疾病；其中，所述医疗专业数据库中包含不同类型疾病与医疗词汇的映射关系；

C、将确定的疾病作为推断出的该医疗文本的潜在疾病进行输出。

2、如权利要求 1 所述的医疗文本的潜在疾病推断方法，其特征在于，所述步骤 A 之前还包括：

从预先确定的数据源获取医疗数据，从所述医疗数据中找出每一种疾病对应的一个或多个医疗词汇，并根据不同类型疾病与医疗词汇的映射关系建立医疗专业数据库。

3、如权利要求 1 所述的医疗文本的潜在疾病推断方法，其特征在于，所述医疗词汇包括：

疾病对应的简介信息、症状信息、并发症信息、治疗药品信息或治疗科室信息中的医疗词汇。

4、如权利要求 1 所述的医疗文本的潜在疾病推断方法，其特征在于，所述医疗专业数据库中还包括疾病对应的各个医疗词汇的权重，所述步骤 B 包括：

基于预先构建的医疗专业数据库，找出该医疗文本中各个医疗词汇对应的疾病，并计算各个疾病对应的医疗词汇的权重加和，选择对应的医疗词汇的权重加和最高的疾病作为确定出的该医疗文本对应的疾病。

5、如权利要求 2 所述的医疗文本的潜在疾病推断方法，其特征在于，所述医疗专业数据库中还包括疾病对应的各个医疗词汇的权重，所述步骤 B 包括：

基于预先构建的医疗专业数据库，找出该医疗文本中各个医疗词汇对应的疾病，并计算各个疾病对应的医疗词汇的权重加和，选择对应的医疗词汇的权重加和最高的疾病作为确定出的该医疗文本对应的疾病。

6、如权利要求 3 所述的医疗文本的潜在疾病推断方法，其特征在于，所述医疗专业数据库中还包括疾病对应的各个医疗词汇的权重，所述步骤 B 包括：

基于预先构建的医疗专业数据库，找出该医疗文本中各个医疗词汇对应的疾病，并计算各个疾病对应的医疗词汇的权重加和，选择对应的医疗词汇的权重加和最高的疾病作为确定出的该医疗文本对应的疾病。

7、如权利要求 1 所述的医疗文本的潜在疾病推断方法，其特征在于，所述对收到的医疗文本进行分词处理的步骤包括：

根据正向最大匹配法将该医疗文本与预先确定的医疗领域专用词汇库进行匹配，得到第一匹配结果，所述第一匹配结果中包含有第一数量的第一词组和第三数量的单字；

根据逆向最大匹配法将该医疗文本与预先确定的医疗领域专用词汇库进行匹配，得到第二匹配结果，所述第二匹配结果中包含有第二数量的第二词组和第四数量的单字；

若所述第一数量与所述第二数量相等，且所述第三数量小于或者等于所述第四数量，则将所述第一匹配结果作为该医疗文本的分词结果；

若所述第一数量与所述第二数量相等，且所述第三数量大于所述第四数量，则将所述第二匹配结果作为该医疗文本的分词结果；

若所述第一数量与所述第二数量不相等，且所述第一数量大于所述第二数量，则将所述第二匹配结果作为该医疗文本的分词结果；

若所述第一数量与所述第二数量不相等，且所述第一数量小于所述第二数量，则将所述第一匹配结果作为该医疗文本的分词结果。

8、如权利要求 2 所述的医疗文本的潜在疾病推断方法，其特征在于，所述对收到的医疗文本进行分词处理的步骤包括：

根据正向最大匹配法将该医疗文本与预先确定的医疗领域专用词汇库进行匹配，得到第一匹配结果，所述第一匹配结果中包含有第一数量的第一词组和第三数量的单字；

根据逆向最大匹配法将该医疗文本与预先确定的医疗领域专用词汇库进行匹配，得到第二匹配结果，所述第二匹配结果中包含有第二数量的第二词组和第四数量的单字；

若所述第一数量与所述第二数量相等，且所述第三数量小于或者等于所述第四数量，则将所述第一匹配结果作为该医疗文本的分词结果；

若所述第一数量与所述第二数量相等，且所述第三数量大于所述第四数量，则将所述第二匹配结果作为该医疗文本的分词结果；

若所述第一数量与所述第二数量不相等，且所述第一数量大于所述第二数量，则将所述第二匹配结果作为该医疗文本的分词结果；

若所述第一数量与所述第二数量不相等，且所述第一数量小于所述第二数量，则将所述第一匹配结果作为该医疗文本的分词结果。

9、如权利要求3所述的医疗文本的潜在疾病推断方法，其特征在于，所述对收到的医疗文本进行分词处理的步骤包括：

根据正向最大匹配法将该医疗文本与预先确定的医疗领域专用词汇库进行匹配，得到第一匹配结果，所述第一匹配结果中包含有第一数量的第一词组和第三数量的单字；

根据逆向最大匹配法将该医疗文本与预先确定的医疗领域专用词汇库进行匹配，得到第二匹配结果，所述第二匹配结果中包含有第二数量的第二词组和第四数量的单字；

若所述第一数量与所述第二数量相等，且所述第三数量小于或者等于所述第四数量，则将所述第一匹配结果作为该医疗文本的分词结果；

若所述第一数量与所述第二数量相等，且所述第三数量大于所述第四数量，则将所述第二匹配结果作为该医疗文本的分词结果；

若所述第一数量与所述第二数量不相等，且所述第一数量大于所述第二数量，则将所述第二匹配结果作为该医疗文本的分词结果；

若所述第一数量与所述第二数量不相等，且所述第一数量小于所述第二数量，则将所述第一匹配结果作为该医疗文本的分词结果。

10、一种电子装置，其特征在于，所述电子装置包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的医疗文本的潜在疾病推断系统，所述医疗文本的潜在疾病推断系统被所述处理器执行时实现如下步骤：

步骤一、对收到的医疗文本进行分词，并将该医疗文本对应的各个分词与预先确定的医疗领域专用词汇库进行匹配，提取出该医疗文本对应的各个分词中的医疗词汇；

步骤二、基于预先构建的医疗专业数据库，确定该医疗文本中的医疗词汇对应的疾病；其中，所述医疗专业数据库中包含不同类型疾病与医疗词汇的映射关系；

步骤三、将确定的疾病作为推断出的该医疗文本的潜在疾病进行输出。

11、如权利要求 10 所述的电子装置，其特征在于，在所述步骤一之前，所述处理器还用于执行所述医疗文本的潜在疾病推断系统，以实现以下步骤：

从预先确定的数据源获取医疗数据，从所述医疗数据中找出每一种疾病对应的一个或多个医疗词汇，并根据不同类型疾病与医疗词汇的映射关系建立医疗专业数据库。

12、如权利要求 10 所述的电子装置，其特征在于，所述医疗词汇包括：

疾病对应的简介信息、症状信息、并发症信息、治疗药品信息或治疗科室信息中的医疗词汇。

13、如权利要求 10 所述的电子装置，其特征在于，所述医疗专业数据库中还包含疾病对应的各个医疗词汇的权重，所述步骤二包括：

基于预先构建的医疗专业数据库，找出该医疗文本中各个医疗词汇对应的疾病，并计算各个疾病对应的医疗词汇的权重加和，选择对应的医疗词汇的权重加和最高的疾病作为确定出的该医疗文本对应的疾病。

14、如权利要求 11 所述的电子装置，其特征在于，所述医疗专业数据库中还包含疾病对应的各个医疗词汇的权重，所述步骤二包括：

基于预先构建的医疗专业数据库，找出该医疗文本中各个医疗词汇对应的疾病，并计算各个疾病对应的医疗词汇的权重加和，选择对应的医疗词汇的权重加和最高的疾病作为确定出的该医疗文本对应的疾病。

15、如权利要求 12 所述的电子装置，其特征在于，所述医疗专业数据库中还包含疾病对应的各个医疗词汇的权重，所述步骤二包括：

基于预先构建的医疗专业数据库，找出该医疗文本中各个医疗词汇对应的疾病，并计算各个疾病对应的医疗词汇的权重加和，选择对应的医疗词汇的权重加和最高的疾病作为确定出的该医疗文本对应的疾病。

16、如权利要求 10 所述的电子装置，其特征在于，所述对收到的医疗文本进行分词处理的步骤包括：

根据正向最大匹配法将该医疗文本与预先确定的医疗领域专用

词汇库进行匹配,得到第一匹配结果,所述第一匹配结果中包含有第一数量的第一词组和第三数量的单字;

根据逆向最大匹配法将该医疗文本与预先确定的医疗领域专用词汇库进行匹配,得到第二匹配结果,所述第二匹配结果中包含有第二数量的第二词组和第四数量的单字;

若所述第一数量与所述第二数量相等,且所述第三数量小于或者等于所述第四数量,则将所述第一匹配结果作为该医疗文本的分词结果;

若所述第一数量与所述第二数量相等,且所述第三数量大于所述第四数量,则将所述第二匹配结果作为该医疗文本的分词结果;

若所述第一数量与所述第二数量不相等,且所述第一数量大于所述第二数量,则将所述第二匹配结果作为该医疗文本的分词结果;

若所述第一数量与所述第二数量不相等,且所述第一数量小于所述第二数量,则将所述第一匹配结果作为该医疗文本的分词结果。

17、如权利要求 11 所述的电子装置,其特征在于,所述对收到的医疗文本进行分词处理的步骤包括:

根据正向最大匹配法将该医疗文本与预先确定的医疗领域专用词汇库进行匹配,得到第一匹配结果,所述第一匹配结果中包含有第一数量的第一词组和第三数量的单字;

根据逆向最大匹配法将该医疗文本与预先确定的医疗领域专用词汇库进行匹配,得到第二匹配结果,所述第二匹配结果中包含有第二数量的第二词组和第四数量的单字;

若所述第一数量与所述第二数量相等,且所述第三数量小于或者等于所述第四数量,则将所述第一匹配结果作为该医疗文本的分词结果;

若所述第一数量与所述第二数量相等,且所述第三数量大于所述第四数量,则将所述第二匹配结果作为该医疗文本的分词结果;

若所述第一数量与所述第二数量不相等,且所述第一数量大于所述第二数量,则将所述第二匹配结果作为该医疗文本的分词结果;

若所述第一数量与所述第二数量不相等,且所述第一数量小于所述第二数量,则将所述第一匹配结果作为该医疗文本的分词结果。

18、如权利要求 12 所述的电子装置,其特征在于,所述对收到的医疗文本进行分词处理的步骤包括:

根据正向最大匹配法将该医疗文本与预先确定的医疗领域专用词汇库进行匹配,得到第一匹配结果,所述第一匹配结果中包含有第一数量的第一词组和第三数量的单字;

根据逆向最大匹配法将该医疗文本与预先确定的医疗领域专用词汇库进行匹配,得到第二匹配结果,所述第二匹配结果中包含有第二数量的第二词组和第四数量的单字;

若所述第一数量与所述第二数量相等,且所述第三数量小于或者等于所述第四数量,则将所述第一匹配结果作为该医疗文本的分词结果;

若所述第一数量与所述第二数量相等,且所述第三数量大于所述第四数量,则将所述第二匹配结果作为该医疗文本的分词结果;

若所述第一数量与所述第二数量不相等,且所述第一数量大于所述第二数量,则将所述第二匹配结果作为该医疗文本的分词结果;

若所述第一数量与所述第二数量不相等,且所述第一数量小于所述第二数量,则将所述第一匹配结果作为该医疗文本的分词结果。

19. 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医疗文本的潜在疾病推断系统,所述医疗文本的潜在疾病推断系统可被至少一个处理器执行,以使所述至少一个处理器执行如下步骤:

对收到的医疗文本进行分词,并将该医疗文本对应的各个分词与预先确定的医疗领域专用词汇库进行匹配,提取出该医疗文本对应的各个分词中的医疗词汇;

基于预先构建的医疗专业数据库,确定该医疗文本中的医疗词汇对应的疾病;其中,所述医疗专业数据库中包含不同类型疾病与医疗词汇的映射关系;

将确定的疾病作为推断出的该医疗文本的潜在疾病进行输出。

20. 如权利要求 19 所述的计算机可读存储介质,其特征在于,在所述对收到的医疗文本进行分词,并将该医疗文本对应的各个分词与预先确定的医疗领域专用词汇库进行匹配,提取出该医疗文本对应的各个分词中的医疗词汇的步骤之前,所述处理器还用于执行所述医疗文本的潜在疾病推断系统,以实现以下步骤:

从预先确定的数据源获取医疗数据,从所述医疗数据中找出每一种疾病对应的一个或多个医疗词汇,并根据不同类型疾病与医疗词汇的映射关系建立医疗专业数据库。

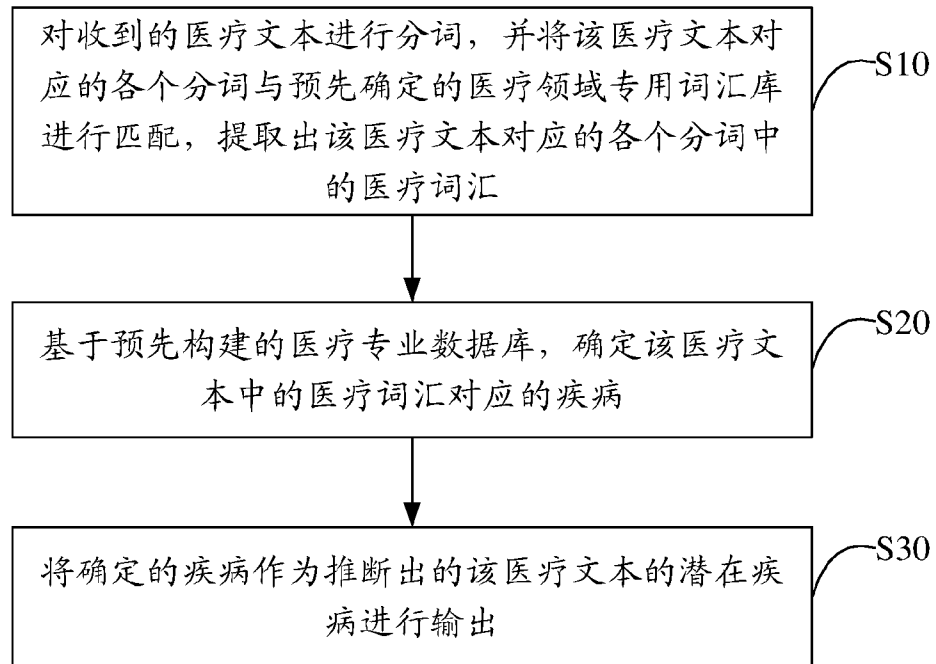


图 1

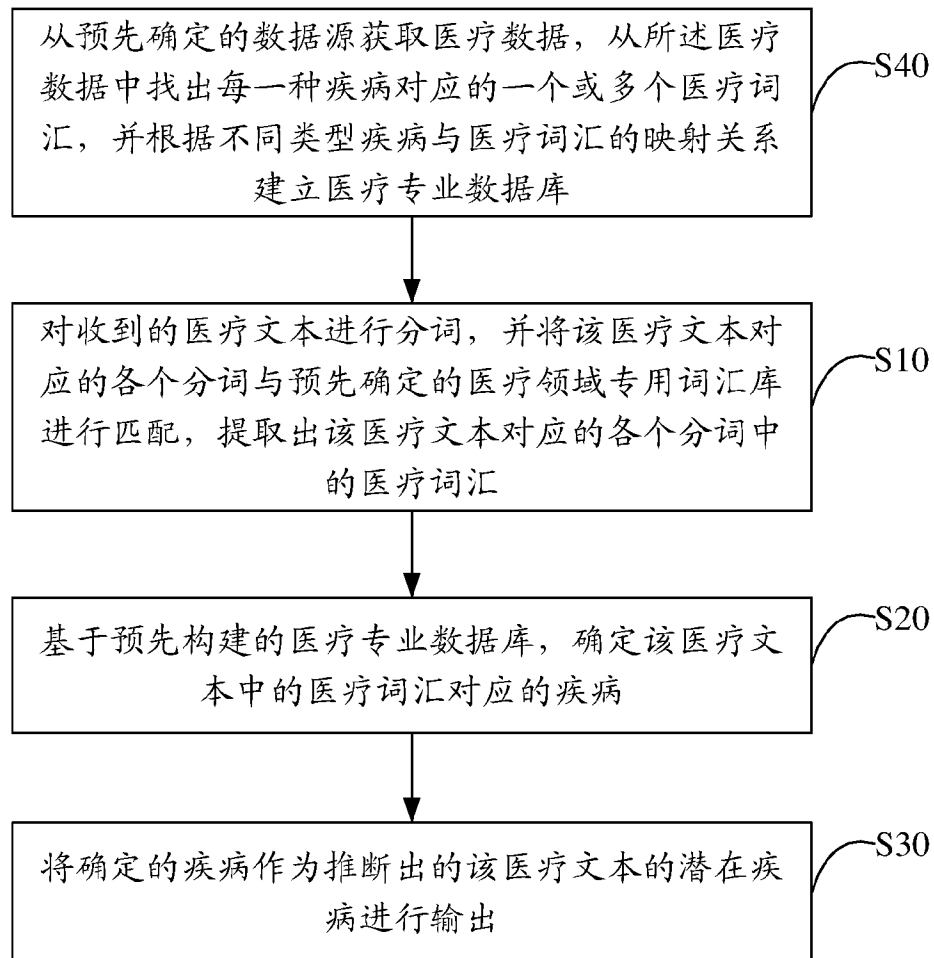


图 2

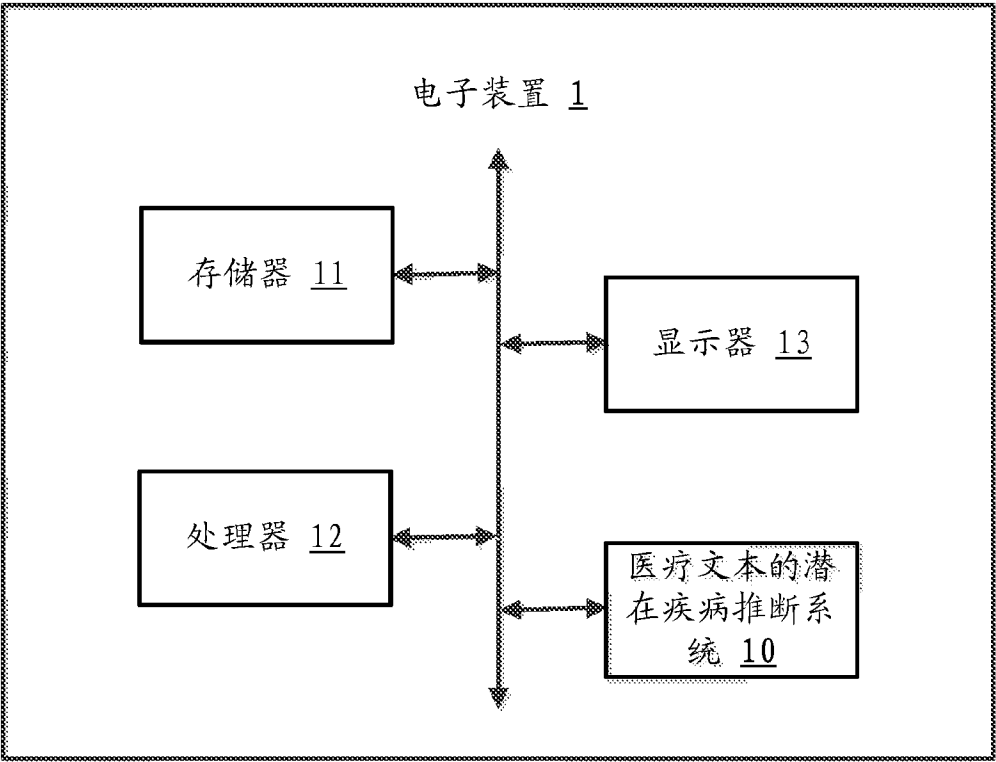


图 3

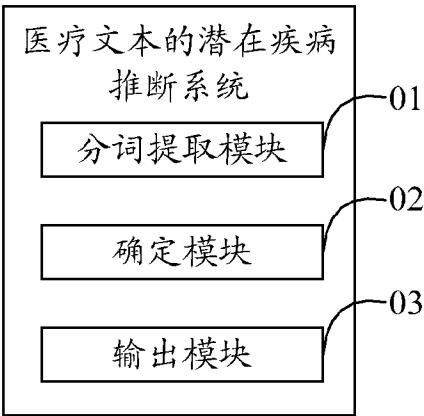


图 4

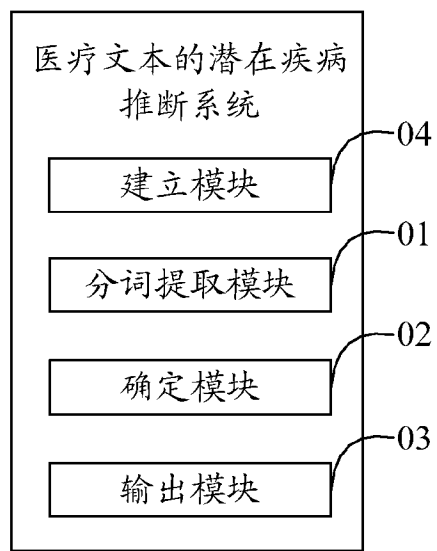


图 5

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/076149

## A. CLASSIFICATION OF SUBJECT MATTER

G06F 19/00 (2018.01) 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)

G06F

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)

CNKI, CNPAT, WPI, EPODOC, IEEE: 医疗, 文本, 疾病, 病人, 数据库, 映射, 诊断, 匹配, medical, text, disease, illness, patient, database, map, diagnostic, match

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                             | Relevant to claim No. |
|-----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| PX        | CN 107680689 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.), 09 February 2018 (09.02.2018), claims 1-10                                                           | 1-20                  |
| X         | CN 106557653 A (HEFEI UNIVERSITY OF TECHNOLOGY), 05 April 2017 (05.04.2017), description, paragraphs 0009-0043 and 0057-0100                                   | 1-20                  |
| X         | CN 105095665 A (EBAONET HEALTHCARE INFORMATION TECHNOLOGY (BEIJING) CO., LTD.), 25 November 2015 (25.11.2015), description, paragraphs 0049-0153 and 0235-0254 | 1-20                  |
| X         | CN 105138829 A (EBAONET HEALTHCARE INFORMATION TECHNOLOGY (BEIJING) CO., LTD.), 09 December 2015 (09.12.2015), description, paragraphs 0063-0233               | 1-20                  |
| A         | CN 106095913 A (GUANGZHOU TONGGOU MEDICAL TECHNOLOGY CO., LTD.), 09 November 2016 (09.11.2016), entire document                                                | 1-20                  |
| A         | US 2010042436 A1 (HAIDER, S.), 18 February 2010 (18.02.2010), entire document                                                                                  | 1-20                  |

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

|                                                                                                                                                                         |                                                                                                                                                                                                                                                  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| * Special categories of cited documents:                                                                                                                                | “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                                              |
| “A” document defining the general state of the art which is not considered to be of particular relevance                                                                | “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                                                                     |
| “E” earlier application or patent but published on or after the international filing date                                                                               | “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 |
| “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) | “&” document member of the same patent family                                                                                                                                                                                                    |
| “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                                                                  |                                                                                                                                                                                                                                                  |

|                                                                                                                                                                                                               |                                                                     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|
| Date of the actual completion of the international search<br>24 April 2018                                                                                                                                    | Date of mailing of the international search report<br>09 May 2018   |
| Name and mailing address of the ISA<br>State Intellectual Property Office of the P. R. China<br>No. 6, Xitucheng Road, Jimenqiao<br>Haidian District, Beijing 100088, China<br>Facsimile No. (86-10) 62019451 | Authorized officer<br>WU, Shouqiu<br>Telephone No. 86-(10)-53961321 |

International application No.  
PCT/CN2018/076149

Form PCT/ISA/210 (patent family annex) (January 2015)

## 国际检索报告

国际申请号

PCT/CN2018/076149

| <b>A. 主题的分类</b><br>G06F 19/00 (2018.01) i<br><br>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>B. 检索领域</b><br>检索的最低限度文献 (标明分类系统和分类号)<br>G06F<br><br>包含在检索领域中的除最低限度文献以外的检索文献<br><br>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))<br>CNKI, CNPAT, WPI, EPODOC, IEEE: 医疗, 文本, 疾病, 病人, 数据库, 映射, 诊断, 匹配, medical, text, disease, illness, patient, database, map, diagnostic, match                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                              |                                              |      |                   |         |    |                                                                      |      |   |                                                                                  |      |   |                                                                                              |      |   |                                                                                   |      |   |                                                                   |      |   |                                                                       |      |
| <b>C. 相关文件</b> <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 107680689 A (平安科技深圳有限公司) 2018年 2月 9日 (2018 - 02 - 09)<br/>权利要求1-10</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 106557653 A (合肥工业大学) 2017年 4月 5日 (2017 - 04 - 05)<br/>说明书第0009-0043、0057-0100段</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 105095665 A (易保互联医疗信息科技北京有限公司) 2015年 11月 25日 (2015 - 11 - 25)<br/>说明书第0049-0153、0235-0254段</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 105138829 A (易保互联医疗信息科技北京有限公司) 2015年 12月 9日 (2015 - 12 - 09)<br/>说明书第0063-0233段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 106095913 A (广州同构医疗科技有限公司) 2016年 11月 9日 (2016 - 11 - 09)<br/>全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 2010042436 A1 (HAIDER, SULTAN) 2010年 2月 18日 (2010 - 02 - 18)<br/>全文</td> <td>1-20</td> </tr> </tbody> </table> |                                                                                              |                                              | 类 型* | 引用文件, 必要时, 指明相关段落 | 相关的权利要求 | PX | CN 107680689 A (平安科技深圳有限公司) 2018年 2月 9日 (2018 - 02 - 09)<br>权利要求1-10 | 1-20 | X | CN 106557653 A (合肥工业大学) 2017年 4月 5日 (2017 - 04 - 05)<br>说明书第0009-0043、0057-0100段 | 1-20 | X | CN 105095665 A (易保互联医疗信息科技北京有限公司) 2015年 11月 25日 (2015 - 11 - 25)<br>说明书第0049-0153、0235-0254段 | 1-20 | X | CN 105138829 A (易保互联医疗信息科技北京有限公司) 2015年 12月 9日 (2015 - 12 - 09)<br>说明书第0063-0233段 | 1-20 | A | CN 106095913 A (广州同构医疗科技有限公司) 2016年 11月 9日 (2016 - 11 - 09)<br>全文 | 1-20 | A | US 2010042436 A1 (HAIDER, SULTAN) 2010年 2月 18日 (2010 - 02 - 18)<br>全文 | 1-20 |
| 类 型*                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 引用文件, 必要时, 指明相关段落                                                                            | 相关的权利要求                                      |      |                   |         |    |                                                                      |      |   |                                                                                  |      |   |                                                                                              |      |   |                                                                                   |      |   |                                                                   |      |   |                                                                       |      |
| PX                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | CN 107680689 A (平安科技深圳有限公司) 2018年 2月 9日 (2018 - 02 - 09)<br>权利要求1-10                         | 1-20                                         |      |                   |         |    |                                                                      |      |   |                                                                                  |      |   |                                                                                              |      |   |                                                                                   |      |   |                                                                   |      |   |                                                                       |      |
| X                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | CN 106557653 A (合肥工业大学) 2017年 4月 5日 (2017 - 04 - 05)<br>说明书第0009-0043、0057-0100段             | 1-20                                         |      |                   |         |    |                                                                      |      |   |                                                                                  |      |   |                                                                                              |      |   |                                                                                   |      |   |                                                                   |      |   |                                                                       |      |
| X                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | CN 105095665 A (易保互联医疗信息科技北京有限公司) 2015年 11月 25日 (2015 - 11 - 25)<br>说明书第0049-0153、0235-0254段 | 1-20                                         |      |                   |         |    |                                                                      |      |   |                                                                                  |      |   |                                                                                              |      |   |                                                                                   |      |   |                                                                   |      |   |                                                                       |      |
| X                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | CN 105138829 A (易保互联医疗信息科技北京有限公司) 2015年 12月 9日 (2015 - 12 - 09)<br>说明书第0063-0233段            | 1-20                                         |      |                   |         |    |                                                                      |      |   |                                                                                  |      |   |                                                                                              |      |   |                                                                                   |      |   |                                                                   |      |   |                                                                       |      |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | CN 106095913 A (广州同构医疗科技有限公司) 2016年 11月 9日 (2016 - 11 - 09)<br>全文                            | 1-20                                         |      |                   |         |    |                                                                      |      |   |                                                                                  |      |   |                                                                                              |      |   |                                                                                   |      |   |                                                                   |      |   |                                                                       |      |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | US 2010042436 A1 (HAIDER, SULTAN) 2010年 2月 18日 (2010 - 02 - 18)<br>全文                        | 1-20                                         |      |                   |         |    |                                                                      |      |   |                                                                                  |      |   |                                                                                              |      |   |                                                                                   |      |   |                                                                   |      |   |                                                                       |      |
| <input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                                              |                                              |      |                   |         |    |                                                                      |      |   |                                                                                  |      |   |                                                                                              |      |   |                                                                                   |      |   |                                                                   |      |   |                                                                       |      |
| * 引用文件的具体类型:<br>“A” 认为不特别相关的表示了现有技术一般状态的文件<br>“E” 在国际申请日的当天或之后公布的在先申请或专利<br>“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)<br>“O” 涉及口头公开、使用、展览或其他方式公开的文件<br>“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<br>“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<br>“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<br>“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<br>“&” 同族专利的文件                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                              |                                              |      |                   |         |    |                                                                      |      |   |                                                                                  |      |   |                                                                                              |      |   |                                                                                   |      |   |                                                                   |      |   |                                                                       |      |
| 国际检索实际完成的日期<br><br>2018年 4月 24日                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                              | 国际检索报告邮寄日期<br><br>2018年 5月 9日                |      |                   |         |    |                                                                      |      |   |                                                                                  |      |   |                                                                                              |      |   |                                                                                   |      |   |                                                                   |      |   |                                                                       |      |
| ISA/CN的名称和邮寄地址<br>中华人民共和国国家知识产权局 (ISA/CN)<br>中国北京市海淀区蓟门桥西土城路6号 100088<br>传真号 (86-10) 62019451                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                              | 授权官员<br><br>武守秋<br><br>电话号码 86-(10)-53961321 |      |                   |         |    |                                                                      |      |   |                                                                                  |      |   |                                                                                              |      |   |                                                                                   |      |   |                                                                   |      |   |                                                                       |      |

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2018/076149

| 检索报告引用的专利文件 |            |    | 公布日<br>(年/月/日) | 同族专利 | 公布日<br>(年/月/日) |
|-------------|------------|----|----------------|------|----------------|
| CN          | 107680689  | A  | 2018年 2月 9日    | 无    |                |
| CN          | 106557653  | A  | 2017年 4月 5日    | 无    |                |
| CN          | 105095665  | A  | 2015年 11月 25日  | 无    |                |
| CN          | 105138829  | A  | 2015年 12月 9日   | 无    |                |
| CN          | 106095913  | A  | 2016年 11月 9日   | 无    |                |
| US          | 2010042436 | A1 | 2010年 2月 18日   | 无    |                |