

(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
2025 年 1 月 9 日 (09.01.2025)



(10) 国际公布号
WO 2025/007940 A1

(51) 国际专利分类号:
G06V 20/20 (2022.01)

(21) 国际申请号: PCT/CN2024/103723

(22) 国际申请日: 2024 年 7 月 4 日 (04.07.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202310814294.0 2023年7月4日 (04.07.2023) CN

(71) 申请人: 阿里巴巴 (中国) 有限公司 (ALIBABA (CHINA) CO., LTD.) [CN/CN]; 中国浙江省杭州市滨江区长河街道网商路 699 号 4 号楼 5 楼 508 室, Zhejiang 310052 (CN)。

(72) 发明人: 张建鹏 (ZHANG, Jianpeng); 中国浙江省杭州市西湖区三墩镇灯彩街 1008 号云谷园区, Zhejiang 310030 (CN)。张灵 (ZHANG, Ling); 美国马里兰州洛克威尔国会巷 210 号 201 室, Maryland 20852 (US)。吕乐 (LU, Le); 美国马里兰州贝塞斯达, Maryland (US)。张剑锋 (ZHANG, Jianfeng); 中国浙江省杭州市余杭区文一西路阿里巴巴西溪园区 969 号, Zhejiang 311121 (CN)。唐禹行 (TANG, Yuxing); 美国马里兰州日耳曼敦车夫路 18105 号, Maryland 20874 (US)。许敏丰 (XU, Minfeng); 中国北京市朝阳区望京东园四区 6 号楼金辉大厦, Beijing 100102 (CN)。郭建飞 (GUO, Jianfei); 中国浙江省杭州市余杭区文一西路阿里巴巴西溪园区 969 号, Zhejiang 311121 (CN)。周靖人 (ZHOU, Jingren); 中国浙江省杭州市西湖区三墩镇灯彩街 1008 号云谷园区, Zhejiang 310030 (CN)。

(54) Title: IMAGE PROCESSING METHOD, ELECTRONIC DEVICE AND STORAGE MEDIUM

(54) 发明名称: 图像处理方法、电子设备及存储介质

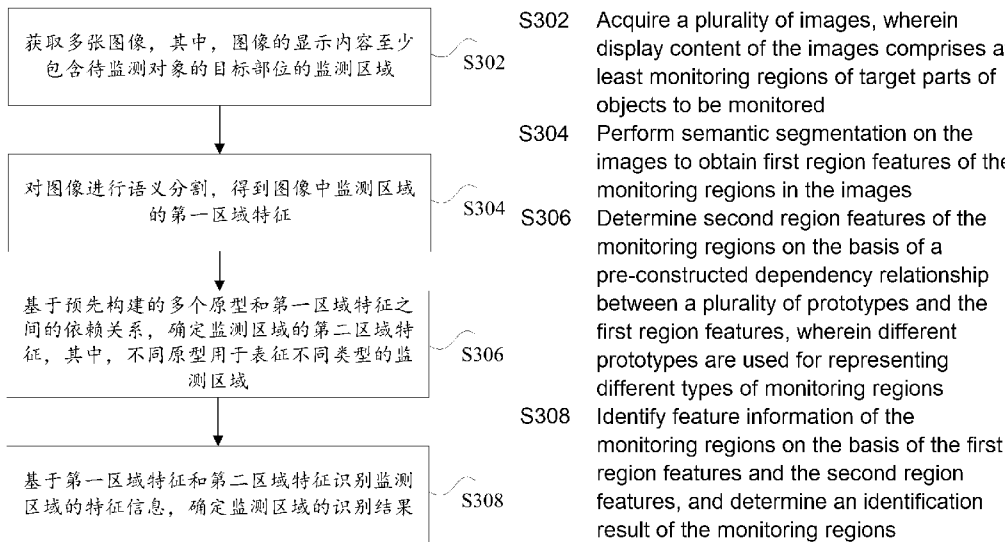


图 3

(57) Abstract: The present application discloses an image processing method, an electronic device and a storage medium. The method is applied to the field of image processing, and comprises: acquiring a plurality of images, wherein display content of the images comprises at least monitoring regions of target parts of objects to be monitored; performing semantic segmentation on the images to obtain first region features of the monitoring regions in the images; determining second region features of the monitoring regions on the basis of a pre-constructed dependency relationship between a plurality of prototypes and the first region features, wherein different prototypes are used for representing different types of monitoring regions; and identifying feature information of the monitoring regions on the basis of the first region features and the second region features, and determining an identification result of the monitoring regions. The

(74) 代理人:北京同钧律师事务所(BEIJING TONGJUN LAW FIRM); 中国北京市海淀区西直门北大街32号院1号楼7层808室, Beijing 100082 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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)。

根据细则4.17的声明:

— 关于申请人有权要求在先申请的优先权(细则4.17(iii))

本国际公布:

— 包括国际检索报告(条约第21条(3))。

present application solves the technical problem in the related art of low identification accuracy of identifying images to be monitored.

(57) 摘要: 本申请公开了一种图像处理方法、电子设备及存储介质。其中, 该方法应用于图像处理领域, 包括: 获取多张图像, 其中, 图像的显示内容至少包含待监测对象的目标部位的监测区域; 对图像进行语义分割, 得到图像中监测区域的第一区域特征; 基于预先构建的多个原型和第一区域特征之间的依赖关系, 确定监测区域的第二区域特征, 其中, 不同原型用于表征不同类型的监测区域; 基于第一区域特征和第二区域特征识别监测区域的特征信息, 确定监测区域的识别结果。本申请解决了相关技术中对待监测图像进行识别的识别准确率低的技术问题。

图像处理方法、电子设备及存储介质

本申请要求于 2023 年 07 月 04 日提交中国专利局、申请号为 202310814294.0、申请
5 名称为“图像处理方法、电子设备及存储介质”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

本申请涉及图像处理领域，具体而言，涉及一种图像处理方法、电子设备及存储介质。

10

背景技术

随着科技的飞速发展，针对图像的处理技术也越来越多地应用到了生活领域中，例如
医学领域和教学领域。现有技术中的图像处理方法，在对图像的目标部位进行处理时，仅
仅只是对图像的目标部位图像进行特征提取，然后对提取到的特征进行识别，得到识别结
15 果，而当图像的清晰度较低时，提取到的特征的准确度就会较低，进而会导致对图像中的
目标部位进行识别的识别准确率低。

针对上述的问题，目前尚未提出有效的解决方案。

发明内容

20 本申请实施例提供了一种图像处理方法、电子设备及存储介质，以至少解决相关技术
中对待监测图像进行识别的识别准确率低的技术问题。

根据本申请实施例的一个方面，提供了一种图像处理方法，包括：获取多张图像，其
中，图像的显示内容至少包含待监测对象的目标部位的监测区域；对图像进行语义分割，
得到图像中监测区域的第一区域特征；基于预先构建的多个原型和第一区域特征之间的依
25 赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；
基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果。

根据本申请实施例的另一方面，还提供了一种图像处理方法，包括：响应作用于操作
界面上的输入指令，在操作界面上显示多张图像，其中，图像的显示内容至少包含待监测
对象的目标部位的监测区域；响应作用于操作界面上的图像处理指令，在操作界面上显示
30 监测区域的识别结果，其中，识别结果是基于第一区域特征和第二区域特征识别监测区域
的特征信息得到的，第二区域特征是基于预先构建的多个原型和第一区域特征之间的依赖
关系确定的，第一区域特征是对医学图像进行语义分割得到的。

根据本申请实施例的另一方面，还提供了一种图像处理方法，包括：在虚拟现实 VR
设备或增强现实 AR 设备的呈现画面上展示多张图像，其中，图像的显示内容至少包含待

监测对象的目标部位的监测区域；对图像进行语义分割，得到图像中监测区域的第一区域特征；基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果；驱动 VR 设备或 AR 设备渲染展示识别结果。

根据本申请实施例的另一方面，还提供了一种图像处理方法，包括：通过调用第一接口获取多张图像，其中，第一接口包括第一参数，第一参数的参数值为多张图像，图像的显示内容至少包含待监测对象的目标部位的监测区域；对图像进行语义分割，得到图像中监测区域的第一区域特征；基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果；通过调用第二接口输出识别结果，其中，第二接口包括第二参数，第二参数的参数值为识别结果。

根据本申请实施例的另一方面，还提供了一种图像处理装置，包括：获取模块，用于获取多张图像，其中，图像的显示内容至少包含待监测对象的目标部位的监测区域；分割模块，用于对图像进行语义分割，得到图像中监测区域的第一区域特征；第一确定模块，用于基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；第二确定模块，用于基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果。

根据本申请实施例的另一方面，还提供了一种图像处理装置，包括：第一显示模块，用于响应作用于操作界面上的输入指令，在操作界面上显示多张图像，其中，图像的显示内容至少包含待监测对象的目标部位的监测区域；第二显示模块，用于响应作用于操作界面上的图像处理指令，在操作界面上显示监测区域的识别结果，其中，识别结果是基于第一区域特征和第二区域特征识别监测区域的特征信息得到的，第二区域特征是基于预先构建的多个原型和第一区域特征之间的依赖关系确定的，第一区域特征是对医学图像进行语义分割得到的。

根据本申请实施例的另一方面，还提供了一种图像处理装置，包括：显示模块，用于在虚拟现实 VR 设备或增强现实 AR 设备的呈现画面上展示多张图像，其中，图像的显示内容至少包含待监测对象的目标部位的监测区域；分割模块，用于对图像进行语义分割，得到图像中监测区域的第一区域特征；第一确定模块，用于基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；第二确定模块，用于基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果；驱动模块，用于驱动 VR 设备或 AR 设备渲染展示识别结果。

根据本申请实施例的另一方面，还提供了一种图像处理装置，包括：获取模块，用于通过调用第一接口获取多张图像，其中，第一接口包括第一参数，第一参数的参数值为多

张图像，图像的显示内容至少包含待监测对象的目标部位的监测区域；分割模块，用于对图像进行语义分割，得到图像中监测区域的第一区域特征；第一确定模块，用于基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；第二确定模块，用于基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果；输出模块，用于通过调用第二接口输出识别结果，其中，第二接口包括第二参数，第二参数的参数值为识别结果。

根据本申请实施例的另一方面，还提供了一种电子设备，包括：存储器，存储有可执行程序；处理器，用于运行程序，其中，程序运行时执行上述任意一项的方法。

根据本申请实施例的另一方面，还提供了一种计算机可读存储介质，计算机可读存储介质包括存储的可执行程序，其中，在可执行程序运行时控制计算机可读存储介质所在设备执行上述任意一项的方法。

根据本申请实施例的另一方面，还提供了一种癌症计算机辅助诊断方法，包括：获取多张医学图像，其中，所述医学图像包含待监测对象的目标监测部位；对所述医学图像进行语义分割，得到所述医学图像中所述监测部位的第一部位特征；基于预先构建的多个原型和所述第一部位特征之间的依赖关系，确定所述监测部位的第二部位特征，其中，不同原型用于表征不同类型的监测部位；基于所述第一部位特征和所述第二部位特征对所述监测部位进行诊断，得到所述监测部位的诊断结果，其中，所述诊断结果用于表征所述监测部位存在恶性病症或良性病症。

根据本申请实施例的另一方面，还提供了一种癌症的计算机辅助诊断系统，包括存储器、处理器以及存储在所述存储器上并在所述处理器上运行的计算机程序，所述处理器执行所述计算机程序可用于执行一种癌症计算机辅助诊断方法，所述方法包括：获取多张医学图像，其中，所述医学图像包含待监测对象的目标监测部位；对所述医学图像进行语义分割，得到所述医学图像中所述监测部位的第一部位特征；基于预先构建的多个原型和所述第一部位特征之间的依赖关系，确定所述监测部位的第二部位特征，其中，不同原型用于表征不同类型的监测部位；基于所述第一部位特征和所述第二部位特征对所述监测部位进行诊断，得到所述监测部位的诊断结果，其中，所述诊断结果用于表征所述监测部位是恶性病症或良性病症。

根据本申请实施例的另一方面，还提供了一种肺癌计算机辅助诊断方法，包括：获取多张医学图像，其中，所述医学图像包含肺部病变；对所述医学图像进行语义分割，得到所述医学图像中肺部病变的第一病变特征；基于预先构建的多个原型和所述第一结节特征之间的依赖关系，确定所述肺部病变的第二病变特征，其中，不同原型用于表征不同类型的肺部病变；基于所述第一病变特征和所述第二病变特征对所述肺结节进行诊断，得到所述肺部病变的诊断结果，其中，所述诊断结果用于表征所述肺部病变是良性病症或恶性病症。

根据本申请实施例的另一方面，还提供了一种肺结节诊断方法，包括：获取多张医学

图像，其中，多张医学图像包含肺结节；对医学图像进行语义分割，得到医学图像中肺结节的第一结节特征；基于预先构建的多个原型和第一结节特征之间的依赖关系，确定肺结节的第二结节特征，其中，不同原型用于表征不同类型的肺结节；基于第一结节特征第二结节特征对肺结节进行诊断，得到肺结节的诊断结果，其中，诊断结果用于表征肺结节是良性结节或恶性结节。

根据本申请实施例的另一方面，还提供了一种肺结节诊断装置，包括：获取模块，用于获取多张医学图像，其中，多张医学图像包含肺结节；分割模块，用于对医学图像进行语义分割，得到医学图像中肺结节的第一结节特征；确定模块，用于基于预先构建的多个原型和第一结节特征之间的依赖关系，确定肺结节的第二结节特征，其中，不同原型用于表征不同类型的肺结节；诊断模块，用于基于第一结节特征第二结节特征对肺结节进行诊断，得到肺结节的诊断结果，其中，诊断结果用于表征肺结节是良性结节或恶性结节。

根据本申请实施例的另一方面，还提供了一种计算机程序，其存储有计算机可执行指令，该指令被处理器执行时实现上述任意一项的方法。

根据本申请实施例的另一方面，还提供了一种计算机程序产品，包括计算机程序，当所述计算机程序在计算机中执行时，令计算机执行上述任意一项的方法。在本申请实施例中，采用获取多张图像；对图像进行语义分割，得到图像中监测区域的第一区域特征；基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征；基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果的方式。容易注意到的是，本申请不仅对图像进行了语义分割，得到区域特征，还结合了原型与区域特征之间的依赖关系，确保最终的区域特征更符合目标部位本身的属性，特征提取准确度更高，进而达到了更准确的对待监测对象的目标部位进行识别的目的，从而实现了提高待监测对象的目标部位的识别准确率的技术效果，进而解决了相关技术中对待监测图像进行识别的识别准确率低的技术问题。

容易注意到的是，上面的通用描述和后面的详细描述仅仅是为了对本申请进行举例和解释，并不构成对本申请的限定。

附图说明

此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

- 图 1 是根据本申请实施例的一种图像处理方法的虚拟现实设备的硬件环境的示意图；
图 2 是根据本申请实施例的一种图像处理方法的计算环境的结构框图；
图 3 是根据本申请实施例 1 的一种图像处理方法的流程图；
图 4 是根据本申请实施例 1 的一种可选的图像处理方法的示意图；
图 5 是根据本申请实施例 2 的一种图像处理方法的流程图；
图 6 是根据本申请实施例 2 的一种可选的操作界面的示意图；

图 7 是根据本申请实施例 3 的一种图像处理方法的流程图;

图 8 是根据本申请实施例 4 的一种图像处理方法的流程图;

图 9 是根据本申请实施例 5 的一种肺结节诊断方法的流程图;

图 10 是根据本申请实施例 5 的一种可选的读者研究与人工智能对比的示意图;

5 图 11 是根据本申请实施例 6 的一种图像处理装置的结构示意图;

图 12 是根据本申请实施例 7 的一种图像处理装置的结构示意图;

图 13 是根据本申请实施例 8 的一种图像处理装置的结构示意图;

图 14 是根据本申请实施例 9 的一种图像处理装置的结构示意图;

图 15 是根据本申请实施例 10 的一种肺结节诊断装置的结构示意图;

10 图 16 是根据本申请实施例的一种计算机终端的结构框图。

具体实施方式

为了使本技术领域的人员更好地理解本申请方案,下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本申
15 请一部分的实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都应当属于本申请保护的范围。

需要说明的是,本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数
20 据在适当情况下可以互换,以便这里描述的本申请的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

首先,在对本申请实施例进行描述的过程中出现的部分名词或术语适用于如下解释:

25 U 型神经网络:包括特征提取网络(编码器),用于特征提取,得到抽象语义特征,以及特征融合网络(解码器),利用前面编码的抽象特征来恢复到原图尺寸的过程,最终得到分割结果(掩码图片),其中,特征提取网络和特征融合网络连接后可以得到形状为 U 型的 U 型神经网络。

30 自注意力模型:查询、键和值来自同一组输入的注意力模型,在处理序列时能更好地理解上下文信息。

交叉注意力模型:键和值相同但与查询不同的注意力模型。

原型:具有相似特征的图像,通过聚类算法将学习过的图像在表征空间上进行聚类,得到的类中心作为该类别的原型。

实施例 1

35 根据本申请实施例,提供了一种图像处理方法,需要说明的是,在附图的流程图示出

的步骤可以在诸如一组计算机可执行指令的计算机系统中执行，并且，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。

图 1 是根据本申请实施例的一种图像处理方法的虚拟现实设备的硬件环境的示意图。如图 1 所示，虚拟现实设备 104 与终端 106 相连接，终端 106 与服务器 102 通过网络进行连接，上述虚拟现实设备 104 并不限于：虚拟现实头盔、虚拟现实眼镜、虚拟现实一体机等，上述终端 104 并不限于 PC、手机、平板电脑等，服务器 102 可以为媒体文件运营商对应的服务器，上述网络包括但不限于：广域网、城域网或局域网。

可选地，该实施例的虚拟现实设备 104 包括：存储器、处理器和传输装置。存储器用于存储应用程序，该应用程序可以用于执行：获取多张图像，其中，图像的显示内容至少包含待监测对象的目标部位的监测区域；对图像进行语义分割，得到图像中监测区域的第一区域特征；基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果，从而解决了相关技术中对待监测图像进行识别的识别准确率低的技术问题，达到了准确对待监测对象的目标部位进行识别的目的。

该实施例的终端可以用于执行在虚拟现实（Virtual Reality，简称为 VR）设备或增强现实（Augmented Reality，简称为 AR）设备的呈现画面上展示多张图像，其中，图像的显示内容至少包含待监测对象的目标部位的监测区域；对图像进行语义分割，得到图像中监测区域的第一区域特征；基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果；驱动 VR 设备或 AR 设备渲染展示识别结果。

可选地，该实施例的虚拟现实设备 104 带有的眼球追踪的 HMD（Head Mount Display，头戴式显示器）头显与眼球追踪模块与上述实施例中的作用相同，也即，HMD 头显中的屏幕，用于显示实时的画面，HMD 中的眼球追踪模块，用于获取用户眼球的实时运动轨迹。该实施例的终端通过跟踪系统获取用户在真实三维空间的位置信息与运动信息，并计算出用户头部在虚拟三维空间中的三维坐标，以及用户在虚拟三维空间中的视野朝向。

图 1 示出的硬件结构框图，不仅可以作为上述 AR/VR 设备（或移动设备）的示例性框图，还可以作为上述服务器的示例性框图，一种可选实施例中，图 2 以框图示出了使用上述图 1 所示的 AR/VR 设备（或移动设备）作为计算环境 201 中计算节点的一种实施例。图 2 是根据本申请实施例的一种图像处理方法的计算环境的结构框图，如图 2 所示，计算环境 201 包括运行在分布式网络上的多个（图中采用 210-1，210-2，...，来示出）计算节点（如服务器）。不同计算节点都包含本地处理和内存资源，终端用户 202 可以在计算环境 201 中远程运行应用程序或存储数据。应用程序可以作为计算环境 201 中的多个服务 220-1，220-2，220-3 和 220-4 进行提供，分别代表服务“A”，“D”，“E”和“H”。

终端用户 202 可以通过客户端上的 web 浏览器或其他软件应用程序提供和访问服务, 在一些实施例中, 可以将终端用户 202 的供应和/或请求提供给入口网关 230。入口网关 230 可以包括一个相应的代理来处理针对服务 (计算环境 201 中提供的一个或多个服务) 的供应和/或请求。

5 服务是根据计算环境 201 支持的各种虚拟化技术来提供或部署的。在一些实施例中, 可以根据基于虚拟机 (Virtual Machine, VM) 的虚拟化、基于容器的虚拟化和/或类似的方式提供服务。基于虚拟机的虚拟化可以通过初始化虚拟机来模拟真实的计算机, 在不直接
10 接触任何实际硬件资源的情况下执行程序 and 应用程序。在虚拟机虚拟化机器的同时, 根据基于容器的虚拟化, 可以启动容器来虚拟化整个操作系统 (Operating System, OS), 以便多个工作负载可以在单个操作系统实例上运行。

在基于容器虚拟化的一个实施例中, 服务的若干容器可以被组装成一个 Pod (例如, Kubernetes Pod)。举例来说, 如图 2 所示, 服务 220-2 可以配备一个或多个 Pod 240-1, 240-2, ..., 240-N (统称为 Pod)。Pod 可以包括代理 245 和一个或多个容器 242-1, 242-2, ..., 242-M (统称为容器)。Pod 中一个或多个容器处理与服务的一个或多个相应功能相关的请
15 求, 代理 245 通常控制与服务相关的网络功能, 如路由、负载均衡等。其他服务也可以配备类似的 Pod。

在操作过程中, 执行来自终端用户 202 的用户请求可能需要调用计算环境 201 中的一个或多个服务, 执行一个服务的一个或多个功能需要调用另一个服务的一个或多个功能。如图 2 所示, 服务“A”220-1 从入口网关 230 接收终端用户 202 的用户请求, 服务“A”220-1
20 可以调用服务“D”220-2, 服务“D”220-2 可以请求服务“E”220-3 执行一个或多个功能。

上述的计算环境可以是云计算环境, 资源的分配由云服务提供上管理, 允许功能的开发无需考虑实现、调整或扩展服务器。该计算环境允许开发人员在不构建或维护复杂基础设施的情况下执行响应事件的代码。服务可以被分割完成一组可以自动独立伸缩的功能, 而不是扩展单个硬件设备来处理潜在的负载。

25 在上述运行环境下, 本申请提供了如图 3 所示的图像处理方法。需要说明的是, 该实施例的图像处理方法可以由图 1 所示实施例的移动终端执行。图 3 是根据本申请实施例 1 的一种图像处理方法的流程图。如图 3 所示, 该方法可以包括如下步骤:

步骤 S302, 获取多张图像, 其中, 图像的显示内容至少包含待监测对象的目标部位的监测区域。

30 上述的待监测对象可以为人体的部位, 但不仅限于此, 还可以是建筑物的局部等。上述的目标部位可以是待监测对象的需要具体监测的部位。例如, 当待监测对象为人体的肺部时, 目标部位可以是肺部的结节、血管、气管等, 但不仅限于此。当待监测对象为建筑物墙上的窗户时, 目标部位可以是窗户的把手、边框角落、中间部位等, 但不仅限于此。上述的监测区域可以是包含有待监测对象的目标部位的区域, 可以称为感兴趣区域 (Region
35 Of Interest, ROI)。

在一种可选的实施例中，当需要对待监测对象的目标部位进行监测时，首先可以获取包含待监测对象的目标部位的原始图像，其次可以基于监测区域对原始图像进行裁剪，得到目标部位的图像，即上述的多张图像，其中，多张图像中至少包含有待监测对象的目标部位。例如，当需要对人体肺部的结节进行监测时，首先可以获取对肺部拍摄得到的 CT 图像，其中，CT 图像为 3 维 (Dimensional) 图像，其次可以基于肺部对 CT 图像进行裁剪得到 3D 裁剪图像，然后为了能够对 3D 裁剪图像进行特征提取，可以将 3D 裁剪图像转换为多张 2D 图像 (即多张图像)。又例如，当需要对建筑物墙上的窗户进行监测时，首先可以获取对窗户进行拍摄得到的多张原始图像，其次可以基于窗户对多张原始图像进行裁剪，即可以得到上述的多张图像。

在另一种可选的实施例中，当需要对待监测对象的目标部位进行监测时，首先可以获取对待监测对象拍摄后得到的视频，其次可以对该视频进行多次抽取得到多张原始图像，然后可以基于监测区域对多张原始图像进行裁剪，得到多张目标部位的图像，即上述的多张图像，其中，多张图像中至少包含有待监测对象的目标部位。例如，当需要对建筑物墙上的窗户进行监测时，首先可以获取对窗户进行拍摄得到视频，其次可以对该视频进行多次抽取得到多张原始图像，然后可以基于窗户对多张原始图像进行裁剪，即可以得到上述的多张图像。

步骤 S304，对图像进行语义分割，得到图像中监测区域的第一区域特征。

在一种可选的实施例中，当获取到图像后，可以通过语义分割模型对监测区域图像进行语义分割，即可以得到监测区域的第一区域特征。例如，可以通过语义分割模型对监测区域图像进行上下文语义分割，即可以得到监测区域的第一区域特征。又例如，首先可以通过语义分割模型对监测区域图像进行上下文语义分割，得到监测区域的预设区域特征，其次还可以通过语义分割模型对预设区域特征进行上下文解析，进而可以得到监测区域的第一区域特征，但不仅限于此。

需要说明的是，上述的语义分割模型可以是相关技术中，任意一种或多种能够对监测区域图像进行语义分割，以得到第一区域特征的模型，在本实施例中不做具体限定。

步骤 S306，基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域。

上述的原型可以是已经确认类型的监测区域，不同原型对应的类型也不同。上述的类型可以是目标部位的类型。例如，当监测区域包含的目标部位为人类的关节时，该监测区域的类型可以是关节，当监测区域包含的目标部位为建筑物的窗户时，该监测区域的类型可以是窗户，但不仅限于此。

在一种可选的实施例中，首先可以构建不同原型与第一区域特征之间的依赖关系，其次当得到第一区域特征后，可以基于依赖关系确定第一区域特征对应的原型，然后可以对第一区域特征以及第一区域特征对应的原型进行处理，得到监测区域的第二区域特征。

步骤 S308，基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区

域的识别结果。

上述的识别结果可以是对监测区域中的目标部位进行识别后得到的结果。例如，当监测区域中的目标部位为关节时，识别结果可以是关节状态好，也可以是关节状态差，当监测区域中的目标部位为窗户时，识别结果可以是窗户符合要求，也可以是窗户不符合要求，
5 但不仅限于此。

在一种可选的实施例中，当得到第一区域特征和第二区域特征后，可以基于第一区域特征和第二区域特征识别监测区域的特征信息，得到监测区域的识别结果。例如，可以分别基于第一区域特征和第二区域特征对监测区域的特征信息进行识别，得到第一识别结果
10 和第二识别结果，然后将第一识别结果和第二识别结果进行对比，选取准确度较高的识别结果作为最终的对比结果。又例如，可以分别基于第一区域特征和第二区域特征对监测区域的特征信息进行识别，得到第一识别结果和第二识别结果，然后取第一识别结果和第二识别结果的平均值，得到最终的识别结果。又例如，还可以首先将第一区域特征和第二区域特征进行特征融合，其次可以基于融合后的区域特征对监测区域的特征信息进行识别，得到监测区域的识别结果，但不仅限于此。

例如，当需要对人类的肺部的结节进行监测时，首先可以获取包含人类肺部的 CT 图像，其次可以基于肺部区域对 CT 图像进行裁剪得到 3D 裁剪图像，然后为了能够对 3D 裁剪图像进行特征提取，可以将 3D 裁剪图像转换为多张 2D 图像（即多张图像）其中，多张图像中均包含有肺部的结节，并通过语义分割模型对肺部图像进行上下文语义分割，得到肺部图像的预设区域特征，然后还可以通过语义分割模型对预设区域特征进行上下文解析，
15 进而可以得到肺部图像的第一区域特征，然后可以基于预先构建的依赖关系确定第一区域特征对应的原型，并对第一区域特征以及第一区域特征对应的原型进行处理得到第二区域特征，最后可以分别基于第一区域特征和第二区域特征对肺部图像的特征信息进行识别，得到第一识别结果和第二识别结果，最后可以获取第一识别结果和第二识别结果的平均值，得到最终的识别结果。

在本申请实施例中，在本申请实施例中，采用获取多张图像；对图像进行语义分割，得到图像中监测区域的第一区域特征；基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征；基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果的方式。容易注意到的是，本申请不仅对图像进行了语义分割，得到区域特征，还结合了原型与区域特征之间的依赖关系，确保最终的区域
20 特征更符合目标部位本身的属性，特征提取准确度更高，进而达到了更准确的对待监测对象的目标部位进行识别的目的，从而实现了提高待监测对象的目标部位的识别准确率的技术效果，进而解决了相关技术中对待监测图像进行识别的识别准确率低的技术问题。

本申请上述实施例中，对图像进行语义分割，得到图像中监测区域的第一区域特征，包括：对图像进行语义分割，得到图像的语义分割结果和全局特征；对语义分割结果和图
30 像进行特征融合，得到融合特征；对全局特征和融合特征进行注意力处理，得到第一区域

特征。

上述的语义分割结果可以是语义掩码 M ，能够体现图像中的像素是否属于目标区域的像素。当像素属于目标区域的像素时，该像素的语义分割结果可以是 1，但不仅如此，还可以是 0。当监测对象为人体肺部时，语义分割结果 M 中包含有不同体素（例如：结节、血管等），属于集合 $\{0: \text{背景}, 1: \text{肺部}, 2: \text{结节}, 3: \text{血管}, 4: \text{气管}\}$ 。

在一种可选的实施例中，首先可以通过语义分割模块对图像进行语义分割，得到图像的语义分割结果和全局特征，其次可以基于目标部位将语义分割结果和图像分割成小块，然后将分割成小块的语义分割结果和图像进行特征融合，也即，可以对包含相同目标部位的，分割成小块后的语义分割结果和图像进行特征融合，得到融合特征，最后可以对全局特征和融合特征进行注意力处理，即可以得到第一区域特征。

本申请上述实施例中，对图像进行语义分割，得到图像的语义分割结果和全局特征，包括：利用 U 型神经网络模型的编码器模块对图像进行特征提取，得到图像的第一图像特征；从 U 型神经网络模型的瓶颈层中，提取全局特征；利用 U 型神经网络模型的编码器模块，对第一图像特征进行解码，得到语义分割结果。

在一种可选的实施例中，首先可以通过 U 型神经网络模型的编码器模块对图像进行特征提取，得到图像的第一图像特征，其次可以通过 U 型神经网络模型的解码器模块对第一图像特征进行解码，即可以得到语义分割结果，此外，还可以通过 U 型神经网络模型的瓶颈层提取图像的全局特征，其中，瓶颈层位于 U 型神经网络模型的中间层。

本申请上述实施例中，对语义分割结果和图像进行特征融合，得到融合特征，包括：分别对语义分割结果和图像进行切分，得到多个子分割结果和多个子图像；分别对多个子分割结果和多个子图像进行特征提取，得到多个子分割结果的子分割特征，以及多个子图像的子图像特征；将子分割特征和子图像特征进行融合，得到融合特征。

在一种可选的实施例中，在得到图像的语义分割结果后，首先可以基于目标部位对语义分割结果和图像进行切分，得到多个子分割结果和多个子图像，其中，多个子分割结果和对应的多个子图像包含有相同的目标部位，其次可以分别对多个子分割结果和多个子图像进行特征提取，得到多个子分割结果的子分割特征和多个子图像的子图像特征，然后将子分割特征和子图像特征进行融合，可以得到融合特征。

本申请上述实施例中，对全局特征和融合特征进行注意力处理，得到第一区域特征，包括：将全局特征和融合特征进行拼接，得到第一拼接特征；利用自注意力模型对第一拼接特征进行自注意力处理，得到第一区域特征。

在一种可选的实施例中，首先可以将全局特征和融合特征进行碎片位置插入（即拼接），即可以得到第一拼接特征 token 为 $[q; t_1, \dots, t_g] \in R^{(g+1)D}$ ，其中， q 为全局特征， t 为融合特征， R 是维度为 $(g+1)D$ 的实数集， D 表示嵌入维度， g 表示融合特征的数量。

在另一种可选的实施例中，首先语义分割结果被切分成小的图像块，和原始图像对应的区域拼接到一起，其次可以通过图像块编码和位置编码，生成一串序列。同时，从卷积

神经网络中抽取出高级语义的特征作为结节的全局特征。

在另一种可选的实施例中，在得到第一拼接特征后，可以通过自注意力模型对第一拼接特征进行自注意力处理，例如，可以通过自注意力模型中的归一化函数（Norm）、自注意力建模（Service Component Architecture, SCA），以及前馈神经网络（Feed Forward Network, FFN）对第一拼接特征进行自注意力处理，即可以得到第一区域特征。

本申请上述实施例中，基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，包括：利用交叉注意力模型对第一区域特征和多个原型进行注意力处理，得到第二区域特征。

在一种可选的实施例中，可以通过交叉注意力模型中的 Norm、交叉原型注意力模块（CPA）以及 FFN 对第一区域特征和多个原型进行注意力处理，即可以得到第二区域特征。

本申请上述实施例中，该方法还包括：获取不同监测区域的全局特征；对不同监测区域的全局特征进行聚类，得到多个特征集合；基于多个特征集合的中心特征，构建多个原型。

在一种可选的实施例中，首先可以获取不同监测区域的全局特征，其次可以对全局特征进行聚类得到多个特征集合 $\{C_1, \dots, C_N\}$ ，其中，C 表示聚类后的特征，N 表示特征数量。

在另一种可选的实施例中，可以通过最小化目标函数 $\sum_{i=1}^N \sum_{p \in C_i} d(p, P_i)$ 以及多个特征集合的中心特征得到多个原型 $P_i = \frac{1}{|C_i|} \sum_{p \in C_i} p$ ，其中，d 为欧几里得函数，p 表示全局特征。需要说明的是，第一原型可以表示为 $P^B \in R^{N/2 \times D}$ ，第二原型可以表示为 $P^M \in R^{N/2 \times D}$ 。

本申请上述实施例中，在基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征之后，该方法还包括：从预先构建的多个原型中，确定与第二区域特征匹配成功的目标原型；对目标原型进行动量更新，得到更新后的区域特征；基于更新后的区域特征对多个原型进行更新。

在一种可选的实施例中，可以通过以下公式对目标原型进行动量更新：

25
$$\begin{cases} P_{\text{argmin}_j d(q, P_j^B)}^B = \lambda \cdot P_{\text{argmin}_j d(q, P_j^B)}^B + (1 - \lambda) \cdot q \\ P_{\text{argmin}_j d(q, P_j^M)}^M = \lambda \cdot P_{\text{argmin}_j d(q, P_j^M)}^M + (1 - \lambda) \cdot q \end{cases}$$
 其中， $P_{\text{argmin}_j d(q, P_j^B)}^B$ 为动量更新后的第一原型， $P_{\text{argmin}_j d(q, P_j^M)}^M$ 为动量更新后的第二原型， λ 为动量因数，一般设置为 0.95，但不仅限于此，其中，动量更新可以帮助加速收敛并提高泛化能力。

在另一种可选的实施例中，当得到更新后的目标原型后，可以基于更新后的目标原型得到更新后的区域特征，然后可以基于更新后的区域特征对多个原型进行更新。

30 本申请上述实施例中，基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果，包括：基于全局特征识别监测区域的特征信息，得到第一子识别结果；基于第一区域特征识别监测区域的特征信息，得到第二子识别结果；基于第二区域特征识别监测区域的特征信息，得到第三子识别结果；对第一子识别结果、第二子识别结果和第三子识别结果进行汇总，得到识别结果。

在一种可选的实施例中，可以通过多层感知器（Multi-Layer Perception, MLP）分别基于全局特征、第一区域特征以及第二区域特征对检测区域的特征信息进行识别，可以分别得到第一识别结果、第二识别结果以及第三识别结果，最后可以将第一识别结果、第二识别结果以及第三识别结果进行汇总，进而可以得到识别结果。例如，可以获取第一识别结果、第二识别结果和第三识别结果的平均值得到识别结果，还可以获取第一识别结果、第二识别结果以及第三识别结果中的较准确的识别结果作为识别结果，但不仅限于此。

图4是根据本申请实施例1的一种可选的图像处理方法的示意图，如图4所示，首先将多张图像输入至U型神经网络模型中，U型神经网络可以对多张图像进行解码和编码后得到语义分割结果，同时U型神经网络模型可以通过瓶颈层输出多张图像的全局特征，其次可以将语义分割结果和多张图像进行切分，得到多个子分割结果和多个子图像，并将多个子分割结果和多个子图像进行特征提取，得到多个子分割特征和子图像特征，然后将多个子分割特征和子图像特征进行融合，可以得到融合特征，如图4中的白色小方块，然后将全局特征和融合特征进行拼接（即块位置嵌入），即可以得到第一拼接特征，如图4中的白色矩形块和斜线阴影矩形块，然后将第一拼接特征输入至自注意力模型中，经过归一化函数、自注意力建模以及前馈神经网络得到第一区域特征，然后将第一区域特征输入至交叉注意力模型中，通过归一化函数、交叉原型注意力以及前馈神经网络得到第二区域特征，最后可以通过MLP从全局特征、第一区域特征以及第二区域特征的特征空间映射到类别空间，得到三个MLP的识别结果，最后获取三个MLP的识别结果的均值，即可以得到最终的识别结果。其中，自注意力模型中的查询（Query）、键（Key）和值（Value）相同，来自同一组输入。而交叉注意力模型中的查询与键和值不同。

需要说明的是，图中的平行四边形中有多个特征集合，其中，箭头连接的多边形为原型，在确定原型后，可以确定与第二区域特征匹配成功的目标原型，然后对目标原型进行动量更新，得到更新后的区域特征；基于更新后的区域特征可以对多个原型进行更新。

需要说明的是，本申请所涉及的用户信息（包括但不限于用户设备信息、用户个人信息等）和数据（包括但不限于用于分析的数据、存储的数据、展示的数据等），均为经用户授权或者经过各方充分授权的信息和数据，并且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准，并提供有相应的操作入口，供用户选择授权或者拒绝。

需要说明的是，对于前述的各方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本申请并不受所描述的动作顺序的限制，因为依据本申请，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作和模块并不一定是本申请所必须的。

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根据上述实施例的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。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，或者网络设备等等）执行本申请各个实施例所述的方法。

5 实施例 2

根据本申请实施例，还提供了一种图像处理方法，需要说明的是，在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行，并且，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。

图 5 是根据本申请实施例 2 的一种图像处理方法的流程图。如图 5 所示，该方法可以包括如下步骤：

步骤 S502，响应作用于操作界面上的输入指令，在操作界面上显示多张图像，其中，图像的显示内容至少包含待监测对象的目标部位的监测区域；

步骤 S504，响应作用于操作界面上的图像处理指令，在操作界面上显示监测区域的识别结果，其中，识别结果是基于第一区域特征和第二区域特征识别监测区域的特征信息得到的，第二区域特征是预先构建的多个原型和第一区域特征之间的依赖关系确定的，第一区域特征是对医学图像进行语义分割得到的。

图 6 是根据本申请实施例 2 的一种可选的操作界面的示意图，如图 6 所示，操作界面包括：输入指令输入区域，处理指令输入区域以及显示区域，当需要对图像中的待监测对象的目标部位进行监测时，首先可以在操作界面的输入指令输入区域输入显示指令，则操作界面可以在显示区域中显示多张显示图像，其次可以在处理指令输入区域中输入处理指令，操作界面可以在显示区域显示监测区域的识别结果，其中，识别结果是基于第一区域特征和第二区域特征识别监测区域的特征信息得到的，第二区域特征是预先构建的多个原型和第一区域特征之间的依赖关系确定的，第一区域特征是对医学图像进行语义分割得到的。

需要说明的是，本申请上述实施例中涉及到的优选实施方案与实施例 1 提供的方案以及应用场景、实施过程相同，但不仅限于实施例 1 所提供的方案。

实施例 3

根据本申请实施例，还提供了一种可以应用于虚拟现实 VR 设备、增强现实 AR 设备等虚拟现实场景下的图像处理方法，需要说明的是，在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行，并且，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。

图 7 是根据本申请实施例 3 的一种图像处理方法的流程图。如图 7 所示，该方法可以包括如下步骤：

步骤 S702，在虚拟现实 VR 设备或增强现实 AR 设备的呈现画面上展示多张图像，其中，图像的显示内容至少包含待监测对象的目标部位的监测区域；

步骤 S704, 对图像进行语义分割, 得到图像中监测区域的第一区域特征;

步骤 S706, 基于预先构建的多个原型和第一区域特征之间的依赖关系, 确定监测区域的第二区域特征, 其中, 不同原型用于表征不同类型的监测区域;

5 步骤 S708, 基于第一区域特征和第二区域特征识别监测区域的特征信息, 确定监测区域的识别结果;

步骤 S7010, 驱动 VR 设备或 AR 设备渲染展示识别结果。

10 在一种可选的实施例中, 当需要对待监测对象的目标部位进行监测时, 首先可以在虚拟现实 VR 设备或增强现实 AR 设备的呈现画面上展示多张图像, 其中, 图像的显示内容至少包含待监测对象的目标部位的监测区域; 其次可以对图像进行语义分割, 得到图像中监测区域的第一区域特征; 然后可以基于预先构建的多个原型和第一区域特征之间的依赖关系, 确定监测区域的第二区域特征, 其中, 不同原型用于表征不同类型的监测区域; 然后可以基于第一区域特征和第二区域特征识别监测区域的特征信息, 确定监测区域的识别结果; 最后可以驱动 VR 设备或 AR 设备渲染展示识别结果。

15 可选地, 在本实施例中, 上述图像处理方法可以应用于由服务器、虚拟现实设备所构成的硬件环境中。在虚拟现实 VR 设备或增强现实 AR 设备的呈现画面上展示识别结果, 服务器可以为媒体文件运营商对应的服务器, 上述网络包括但不限于: 广域网、城域网或局域网, 上述虚拟现实设备并不限于: 虚拟现实头盔、虚拟现实眼镜、虚拟现实一体机等。

20 可选地, 虚拟现实设备包括: 存储器、处理器和传输装置。存储器用于存储应用程序, 该应用程序可以用于执行: 在虚拟现实 VR 设备或增强现实 AR 设备的呈现画面上展示多张图像, 其中, 图像的显示内容至少包含待监测对象的目标部位的监测区域; 对图像进行语义分割, 得到图像中监测区域的第一区域特征; 基于预先构建的多个原型和第一区域特征之间的依赖关系, 确定监测区域的第二区域特征, 其中, 不同原型用于表征不同类型的监测区域; 基于第一区域特征和第二区域特征识别监测区域的特征信息, 确定监测区域的识别结果; 驱动 VR 设备或 AR 设备渲染展示识别结果。

需要说明的是, 该实施例的上述应用在 VR 设备或 AR 设备中的图像处理方法可以包括图 3 所示实施例的方法, 以实现驱动 VR 设备或 AR 设备展示识别结果的目的。

30 可选地, 该实施例的处理器可以通过传输装置调用上述存储器存储的应用程序以执行上述步骤。传输装置可以通过网络接收服务器发送的媒体文件, 也可以用于上述处理器与存储器之间的数据传输。

35 可选地, 在虚拟现实设备中, 带有眼球追踪的头戴式显示器, 该 HMD 头显中的屏幕, 用于显示展示的视频画面, HMD 中的眼球追踪模块, 用于获取用户眼球的实时运动轨迹, 跟踪系统, 用于追踪用户在真实三维空间的位置信息与运动信息, 计算处理单元, 用于从跟踪系统中获取用户的实时位置与运动信息, 并计算出用户头部在虚拟三维空间中的三维坐标, 以及用户在虚拟三维空间中的视野朝向等。

在本申请实施例中，虚拟现实设备可以与终端相连接，终端与服务器通过网络进行连接，上述虚拟现实设备并不限于：虚拟现实头盔、虚拟现实眼镜、虚拟现实一体机等，上述终端并不限于 PC、手机、平板电脑等，服务器可以为媒体文件运营商对应的服务器，上述网络包括但不限于：广域网、城域网或局域网。

5 需要说明的是，本申请上述实施例中涉及到的优选实施方案与实施例 1 提供的方案以及应用场景、实施过程相同，但不仅限于实施例 1 所提供的方案。

实施例 4

根据本申请实施例，还提供了一种图像处理方法，需要说明的是，在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行，并且，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。

10 图 8 是根据本申请实施例 4 的一种图像处理方法的流程图。如图 8 所示，该方法可以包括如下步骤：

步骤 S802，通过调用第一接口获取多张图像，其中，第一接口包括第一参数，第一参数的参数值为多张图像，图像的显示内容至少包含待监测对象的目标部位的监测区域；

15 步骤 S804，对图像进行语义分割，得到图像中监测区域的第一区域特征；

步骤 S806，基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；

步骤 S808，基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果；

20 步骤 S8010，通过调用第二接口输出识别结果，其中，第二接口包括第二参数，第二参数的参数值为识别结果。

上述的第一接口可以是向服务器获取多张图像的接口，上述的第二接口可以是向服务器发送识别结果的接口。

在一种可选的实施例中，当需要对待监测对象的目标部位进行监测时，首先可以通过
25 调用第一接口获取多张图像，其中，第一接口包括第一参数，第一参数的参数值为多张图像，图像的显示内容至少包含待监测对象的目标部位的监测区域；其次可以对图像进行语义分割，得到图像中监测区域的第一区域特征；然后可以基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；然后可以基于第一区域特征和第二区域特征识别监测区域的特征信息，
30 确定监测区域的识别结果；最后可以通过调用第二接口输出识别结果，其中，第二接口包括第二参数，第二参数的参数值为识别结果。

需要说明的是，本申请上述实施例中涉及到的优选实施方案与实施例 1 提供的方案以及应用场景、实施过程相同，但不仅限于实施例 1 所提供的方案。

实施例 5

35 根据本申请实施例，还提供了一种肺结节诊断方法，需要说明的是，在附图的流程图

示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行，并且，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。

图 9 是根据本申请实施例 5 的一种肺结节诊断方法的流程图。如图 9 所示，该方法可以包括如下步骤：

步骤 S902，获取多张医学图像，其中，多张医学图像包含肺结节。

上述的多张医学图像可以是多张 2D 医学图像，是通过对人体的电子计算机断层摄影（computer tomography, CT）图片的 ROI 区域（例如可以是肺部区域）进行裁剪并进行转换后得到的多张 2D 图像，其中，多张医学图像中包含有肺结节。

在一种可选的实施例中，当需要对肺部进行癌症筛查时，首先可以获取包含肺部结节的医学图像，其次可以基于肺部区域对医学图像进行裁剪得到肺部图像，最后可以将肺部图像进行转换得到多张 2D 医学图像，其中，多张医学图像中包含有肺结节。

步骤 S904，对医学图像进行语义分割，得到医学图像中肺结节的第一结节特征；

在一种可选的实施例中，当获取到医学图像后，首先可以通过语义分割模型对肺部图像进行语义分割，即可以得到肺结节的第一结节特征。例如，可以通过语义分割模型对肺部图像进行上下文语义分割，即可以得到肺部图像中肺结节的第一结节特征。又例如，首先可以通过语义分割模型对肺部图像进行上下文语义分割，得到肺部图像中肺结节的第一预设结节特征，其次还可以通过语义分割模型对第一预设结节特征进行上下文解析，进而可以得到肺部图像中肺结节的第一结节特征，但不仅限于此。

例如，在上下文分割阶段，结节上下文信息对良性和恶性诊断具有重要影响。例如，与输血管相关的结节比孤立的结节更可能是恶性的。因此，可以使用 U 型神经网络（UNet）解析输入图像（即医学图像）的语义掩码 m （即上下文语义分割图像），其中，输入图像是对原始 CT 图像中肺结节 ROI 区域进行裁剪得到的图像，输入是一个三维的 volume，是由多张 2D 图像（slice）组成。从而允许随后对结节及其周围结构两者进行上下文建模。具体地， m 的每个体素属于 {0: 背景, 1: 肺, 2: 结节, 3: 血管, 4: 气管}。这种分割过程可以收集对于准确诊断至关重要的综合上下文信息。为了诊断的目的，可以从 UNet 的瓶颈中提取全局特征作为结节嵌入 q ，这将在后面的诊断阶段使用。

需要说明的是，在上下文分割阶段，肺结节良恶性鉴别时需要的上下文内容包括肺部正常组织、结节、血管和气管，这些信息不仅反映了结节本身的形状、位置、和大小等信息，同时也反映了结节和周围组织的结构的关系。本申请使用 U 型卷积神经网络对上述的上下文语义信息进行像素级别的识别，可以终得到每一个结节的上下文语义分割图。

步骤 S906，基于预先构建的多个原型和第一结节特征之间的依赖关系，确定肺结节的第二结节特征，其中，不同原型用于表征不同类型的肺结节。

在一种可实施例中，首先可以构建不同原型与第一结节特征之间的依赖关系，其次当得到第一结节特征后，可以基于依赖关系确定第一结节特征对应的原型，然后可以对第一

结节特征以及第一结节特征对应的原型进行处理，得到肺部图像的肺结节的第二结节特征。

例如，在结节内部上下文解析阶段，本申请设计了一种基于注意力机制的上下文解析模块，用来深入剖析结节，整合其上下文信息，提升良恶性鉴别能力。具体地，针对不同结节，其上下文语义分割图像被切分成小的图像块，和原始图像（即输入图像）对应的区域拼接到一起得到多个重叠块，通过一个图像块编码和位置编码，生成一串上下文特征（token）。同时，从卷积神经网络中抽取出高级语义的特征作为结节的全局表征，也称之为结节 token。通过设计一种上下文自注意力模块，建模结节 token 和上下文 token 之间的长距离的依赖关系，从上下文信息中提取相关的良恶性鉴别依据。自注意力模块输出的结节 token 作为结节良恶性鉴别的新表征。

需要说明的是，在结节内部上下文解析阶段，可以通过聚合由分割模型产生的上下文信息来增强结节的区别表示。具体地，可以通过重叠块嵌入，将上下文掩码符号化到一组序列中。输入图像也被分割成小块，并被嵌入上下文 tokens 中以保持原始图像信息。此外，以可学习的方式添加位置编码以保持位置信息。可以将结节嵌入 token 预先附加到上下文序列，表示为 $[q; t_1, \dots, t_g] \in \mathbb{R}^{(g+1)D}$ 。其中， g 是上下文 tokens 的数量， D 表示嵌入维度。然后可以同时对这些 tokens 进行自注意力建模，称为 SCA，用于将上下文信息聚合到结节嵌入中。在最后一个 SCA 块的输出处的结节嵌入 token 用作更新的结节表示。对结节嵌入及其背景结构之间的依赖性进行明确建模可以导致更区别的表示的进化，从而改善良性和恶性节瘤之间的区别。

又例如，在结节原型召回学习阶段，本申请设计了一种结节诊断知识原型回顾模块，首先结节的原型定义为具有相似特征的结节代表，通过聚类算法将学习过的肺结节在表征空间上进行聚类，得到的类中心作为该类别的原型。其中，原型具有良恶性区分，良性的原型是通过相似特征的良性结节计算得到，而恶性的原型则是来自于恶性结节。为了利用这些原型知识，本申请设计了一种交叉原型注意力模块构建当前结节和其他原型之间的关系，其中 query 是来自当前结节的表征，key 和 value 分别是来自原型的表征。该交叉注意力模块输出的 query token 作为最终的良恶性鉴别表征。

需要说明的是，为了保留先前获取的知识，需要更有效的方法，而不是将所有学习的结节存储在存储器中，这将导致存储和计算资源的浪费。为了简化这个过程，可以将相关结节凝聚成原型的形式。对于一组结节（即多张结节图像），可以将其聚集成 N 个组 $\{C_1, \dots, C_N\}$ ，通过最小化目标函数 $\sum_{i=1}^N \sum_{p \in C_i} d(p, P_i)$ ，其中， d 是欧几里得距离函数， p 表示结节嵌入，并将每个集群的中心作为原型。考虑到良性和恶性结节之间的差异，可以将原型分为良性组和恶性组，由 $P^B \in \mathbb{R}^{N/2 \times D}$ 和 $P^M \in \mathbb{R}^{N/2 \times D}$ 表示。其中，除了解析内部上下文，还可以捕捉结节和外部原型之间的层间依赖性。这使得 PARE 能够探索除单个结节外的相关鉴定基础。为了实现这一点，本申请设计了交叉原型注意力（CPA）模块，其利用结节嵌入作为查询以及原型作为关键字和值。其允许结节嵌入选择性地参与原型序列的最相关部分。在最后的 CPA 模块的输出处的查询的状态作为最终结节表示以预测其恶性标记，

“良性” ($y=0$) 或 “恶性” ($y=1$)。其中, PARE 为本申请提出的一种对肺结节进行诊断的模型, 该模型包括: 上下文分割、结节内部上下文解析以及结节原型召回学习三部分。

需要说明的是, 本申请还可以在线方式更新原型, 从而允许原型快速调整结节嵌入的变化。对于数据为 (x, y) 的结节嵌入 q , 选出其最近的原型, 然后通过以下动量规则更新:

$$\begin{cases} P_{\text{argmin}_j d(q, P_j^B)}^B = \lambda \cdot P_{\text{argmin}_j d(q, P_j^B)}^B + (1 - \lambda) \cdot q, & \text{if } y = 0 \\ P_{\text{argmin}_j d(q, P_j^M)}^M = \lambda \cdot P_{\text{argmin}_j d(q, P_j^M)}^M + (1 - \lambda) \cdot q, & \text{otherwise} \end{cases}$$

其中, $P_{\text{argmin}_j d(q, P_j^B)}^B$ 为动量更新后的良性原型, $P_{\text{argmin}_j d(q, P_j^M)}^M$ 为动量更新后恶性原型, λ 为动量因数, 一般设置为 0.95, 但不仅限于此, 其中, 动量更新可以帮助加速收敛并提高泛化能力。

10 步骤 S908, 基于第一结节特征第二结节特征对肺结节进行诊断, 得到肺结节的诊断结果, 其中, 诊断结果用于表征肺结节是良性结节或恶性结节。

在一种可选的实施例中, 当得到第一结节特征和第二结节特征后, 可以基于第一结节特征和第二结节特征识别对肺结节进行诊断, 得到肺结节的诊断结果。例如, 可以分别基于第一结节特征和第二结节特征对肺结节进行诊断, 得到第一诊断结果和第二诊断结果, 15 然后将第一诊断结果和第二诊断结果进行对比, 选取准确度较高的诊断结果作为最终的诊断结果。又例如, 可以分别基于第一结节特征和第二结节特征对对肺结节进行诊断, 得到第一诊断结果和第二诊断结果, 然后取第一诊断结果和第二诊断结果的平均值, 得到最终的诊断结果。又例如, 还可以首先将第一结节特征和第二结节特征进行特征融合, 其次可以基于融合后的结节特征对对肺结节进行诊断, 得到肺结节的诊断结果, 但不仅限于此。

20 需要说明的是, 本申请设计了一种深度监督训练模式来提高良恶性鉴别能力。深度监督信号分别在卷积神经网络输出的结节全局表征、自上下文注意力模块输出的结节表征、以及交叉原型注意力模块输出的结节表征。通过添加一个多层感知器 (Multi-Layer Perception, MLP), 从各自的表征空间映射到良性和恶性两大类空间。在推理场景下, 三个 MLP 得到良恶性类别概率通过平均的方式集成为最终的鉴别概率。

25 本申请提出一种放射科医生激励方法, 模拟放射科医生的诊断过程, 由上下文解析和原型回顾模块组成。上下文解析模块首先对结节的上下文结构进行分段, 然后聚合上下文信息, 以更全面地理解结节。原型回顾模块利用基于原型的学习来将先前学习的情况压缩为用于比较性分析的原型, 其在训练期间以动量方式在线更新。基于这两个模块, 本申请的方法利用节结的固有特性和从其他节结积累的外部知识两者来实现合理的诊断。为了满 30 足低剂量和非连续筛选的需要, 分别从低剂量和非连续 CT 收集 12852 和 4029 个结节的大规模数据集, 每个具有病理学或后续确认的标记。在几个数据集上的实验证明, 本申请的方法在低剂量和不连续两种场景上实现了先进的筛选性能。

本申请上述实施例中: 通过预先构建的多个原型和第一结节之间的依赖关系, 能够从结节及其周围器官组织中提取和聚合丰富的上下文信息; 通过预先将学习到的结节诊断知

识压缩成原型，并将其用作参考，协助诊断新的结节，能够最终的区域特征更符合目标部位本身的属性，特征提取准确度更高；基于第一结节特征和第二结节特征进行诊断，能够实现面向低剂量和平扫两大筛查场景的肺结节良恶性鉴别，提高了临床应用的通用性。

表 1 是根据本申请实施例 5 的一种可选的超参数的消融比较，在表 1 中，本申请调查了不同配置对 PARE 性能对验证集的影响，包括变压器层、原型数量、嵌入维数和深度监督。由表 1 可知，可以通过增加变换器层的数量、增加原型的数量、使 token 嵌入的通道尺寸加倍、或使用深度分类监督来获得更高的 AUC 评分。基于 0.931 的最高 AUC 评分，在以下实验中经验地设定 L=4、N=40、D=256 和 DS=真。其中，超参数包括：变压器层（L）、原型数量（N）、嵌入维度（D）以及深度监督（DS）。

表 1 超参数的消融比较

L	N	D	DS	AUC
1	20	128	√	0.912
2	20	128	√	0.918
4	20	128	√	0.924
4	10	128	√	0.920
4	40	128	√	0.924
4	40	256	√	0.931
4	40	256	×	0.926

表 2 是根据本申请实施例 5 的一种可选的不同模块的有效性，在表 2 中，本申请调查了不同方法/模块在验证集合上的消融研究并且观察以下结果：（1）纯分割方法表现得比纯分类方法更好，主要是因为它使得能够在像素级上进行更大的监督，（2）联合分割和分类优于任何单一方法，表明两个任务的互补效果，（3）上下文解析和原型比较两者有助于在强基线上改善性能，从而证明两个模块的有效性，并且（4）与仅仅分割结节相比，分割更多上下文结构（例如血管、肺和气管）提供了轻微的改进。其中，MT 表示多任务学习。上下文：帧内上下文解析。原型：原型间的回顾。*表示在分割任务中仅使用结节掩模。

表 2 不同模块的有效性

Method	AUC
Pure classification	0.907
Pure segmentation	0.915
MT	0.916
MT+Context*	0.921
MT+Context	0.924
MT+Context+Prototype	0.931

在两个筛选场景上与其他方法的比较：表 3 是根据本申请实施例 5 的一种可选的 NNLST 和内部测试集上不同方法的比较，包括基于纯分类的方法、基于纯分割的方法和基于多任务的方法。基于结节尺寸分布，在两个测试组中进行分层评估。这些结果指示该基于分段的方法胜过纯粹的分类方法，主要是由于其对上下文结构进行分段的优越能力。此

外，基于多任务的 CA-网胜过任何单任务方法。在 NLST 和内部测试集上，本申请的 PARE 方法超过了大多数其他方法。此外，通过利用多个深度监督头的集合，对两个数据集将总体 AUC 进一步改善至 0.931。其中，。†表示纯分类；‡表示纯分割；◊表示多任务学习；*表示深度监督头集合。需要注意的是：本申请在 CA-Net 中加入了分割任务。

5 表 3 NNLST 和内部测试集上不同方法的比较

Method	NLST test set				In-house test set			
	< 10mm	10~20mm	> 20mm	All	< 10mm	10~20mm	> 20mm	All
CNN†	0.742	0.706	0.780	0.894	0.851	0.797	0.744	0.901
ASPP[5]†	0.798	0.716	0.801	0.902	0.854	0.788	0.743	0.901
MiT[24]†	0.821	0.755	0.810	0.908	0.858	0.784	0.751	0.904
nn Unet[8]‡	0.815	0.736	0.815	0.910	0.863	0.804	0.750	0.911
CA-Net[12]◊	0.833	0.759	0.807	0.916	0.878	0.786	0.779	0.918
PARE◊	0.882	0.770	0.826	0.928	0.892	0.817	0.783	0.927
PARE◊*	0.890	0.781	0.827	0.931	0.899	0.821	0.780	0.931

LUNGx 的外部评价：本申请使用 LUNGx 作为外部检验来评价的 PARE 的泛化。值得注意的是，这些比较的方法从未在 LUNGx 上进行过训练。表 4 是根据本申请实施例 5 的一种可选的与其他方法在 LUNGx 上的对比，从表 4 可以看出，本申请的 PARE 模型的 AUC 最高为 0.801，比方法 DAR 提高了 2%。本申请还进行了一项读者研究，将 PARE 与两位
10 分别具有 8 年和 13 年肺结节诊断经验的具有丰富经验的放射科医生进行比较。图 10 是根据本申请实施例 5 的一种可选的读者研究与人工智能对比的示意图，图 3 的结果显示，本申请的方法达到了与放射科医生相当的性能。

表 4 与其他方法在 LUNGx 上的对比

Method	AUC
NLNL[9]	0.683
CIRDataset[6]	0.743
D2CNN[25]	0.746
KBC[23]	0.768
DAR[11]	0.781
PARE (Ours)	0.801

需要说明的是，本申请所涉及的用户信息（包括但不限于用户设备信息、用户个人信息等）和数据（包括但不限于用于分析的数据、存储的数据、展示的数据等），均为经用户
15 授权或者经过各方充分授权的信息和数据，并且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准，并提供有相应的操作入口，供用户选择授权或者拒绝。

LDCT 和 NCCT 的泛化：本申请的模型实在 LDCT 和 NCCT 数据集的混合上训练的，
20 可以在低剂量和常规剂量应用中表现良好。表 5 示出了在三种训练配置下获得的模型的泛化性能比较，从表 5 可以看出，单独在 LDCT 或 NCCT 数据集上训练的模型不能很好地推广到其他模态，至少有 6%的 AUC 下降。然而，本申请的混合训练方法在 LDCT 和 NCCT

上都表现最好，几乎没有性能下降。

表 5 不同训练配置对 LDCT 和 NCCT 性能的影响对比

训练场景	测试场景		两个场景之间的性能差距
	LDCT	NCCT	
LDCT	0.927	0.867	6.0%
NCCT	0.802	0.893	9.1%
LDCT+NCCT	0.928	0.927	0.1%

此外通过比较由医生和其他由 TotalSegmentator 生成的语义掩码标注的从 LIDC-IDRI 数据中得到的分割结果，和由本申请实施例提供的在分割任务中使用的结节掩模标注的分割结果，PARE 在 LIDC-IDRI 的 2630 个结节上达到 77.9% 的 Dice 得分。

需要说明的是，本申请上述实施例中涉及到的优选实施方案与实施例 1 提供的方案以及应用场景、实施过程相同，但不仅限于实施例 1 所提供的方案。

实施例 6

根据本申请实施例，还提供了一种癌症计算机辅助诊断方法，该方法包括：
获取多张医学图像，其中，所述医学图像包含待监测对象的目标监测部位；
对所述医学图像进行语义分割，得到所述医学图像中所述监测部位的第一部位特征；
基于预先构建的多个原型和所述第一部位特征之间的依赖关系，确定所述监测部位的第二部位特征，其中，不同原型用于表征不同类型的监测部位；
基于所述第一部位特征和所述第二部位特征对所述监测部位进行诊断，得到所述监测部位的诊断结果，其中，所述诊断结果用于表征所述监测部位是恶性病症或良性病症。

需要说明的是，本申请上述实施例中涉及到的优选实施方案与实施例 1 提供的方案以及应用场景、实施过程相同，但不仅限于实施例 1 所提供的方案。

实施例 7

根据本申请实施例，还提供了一种癌症的计算机辅助诊断系统，包括存储器、处理器以及存储在所述存储器上并在所述处理器上运行的计算机程序，所述处理器执行所述计算机程序可用于执行一种癌症计算机辅助方法，所述方法包括：

获取多张医学图像，其中，所述医学图像包含待监测对象的目标监测部位；
对所述医学图像进行语义分割，得到所述医学图像中所述监测部位的第一部位特征；

基于预先构建的多个原型和所述第一部位特征之间的依赖关系，确定所述监测部位的第二部位特征，其中，不同原型用于表征不同类型的监测部位；

基于所述第一部位特征和所述第二部位特征对所述监测部位进行诊断，得到所述监测部位的诊断结果，其中，所述诊断结果用于表征所述监测部位是恶性病症或良性病症。

需要说明的是，本申请上述实施例中涉及到的优选实施方案与实施例 1 提供的方案以及应用场景、实施过程相同，但不仅限于实施例 1 所提供的方案。

实施例 8

根据本申请实施例的另一方面，还提供了一种肺癌计算机辅助诊断方法，包括：

获取多张医学图像，其中，所述医学图像包含肺部病变；

对所述医学图像进行语义分割，得到所述医学图像中肺部病变的第一病变特征；

5 基于预先构建的多个原型和所述第一结节特征之间的依赖关系，确定所述肺部病变的第二病变特征，其中，不同原型用于表征不同类型的肺部病变；

基于所述第一病变特征和所述第二病变特征对所述肺结节进行诊断，得到所述肺部病变的诊断结果，其中，所述诊断结果用于表征所述肺部病变是良性病症或恶性病症。

10 在本实施例中，肺部病变可以包括肺结节。若诊断结果为肺部病变为恶性病症，则该肺部病变可能是由肺癌造成，以便医生或诊断设备根据诊断结果确定治疗方案。

实施例 9

根据本申请实施例，还提供了一种用于实施上述图像处理方法的图像处理装置。图 11 是根据本申请实施例 6 的一种图像处理装置的结构示意图，如图 11 所示，该装置包括：获取模块 1102、分割模块 1104、第一确定模块 1106 和第二确定模块 1108。

15 其中，获取模块用于获取多张图像，其中，图像的显示内容至少包含待监测对象的目标部位的监测区域；分割模块用于对图像进行语义分割，得到图像中监测区域的第一区域特征；第一确定模块用于基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；第二确定模块用于基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果。

20 此处需要说明的是，上述获取模块、分割模块、第一确定模块和第二确定模块对应于实施例 1 中的步骤 S302 至步骤 S308，四个模块与对应的步骤所实现的实例和应用场景相同，但不限于上述实施例 1 所公开的内容。需要说明的是，上述模块或单元可以是存储在存储器中并由一个或多个处理器处理的硬件组件或软件组件，上述模块也可以作为装置的一部分可以运行在实施例 1 提供的 AR/VR 设备中。

本申请上述实施例中，分割模块包括：分割单元、融合单元和第一处理单元。

其中，分割单元用于对图像进行语义分割，得到图像的语义分割结果和全局特征；融合单元，用于对语义分割结果和图像进行特征融合，得到融合特征；第一处理单元用于对全局特征和融合特征进行注意力处理，得到第一区域特征。

30 本申请上述实施例中，分割单元包括：第一提取子单元、第二提取子单元和解码子单元。

其中，第一提取子单元用于利用 U 型神经网络模型的编码器模块对图像进行特征提取，得到图像的第一图像特征；第二提取子单元用于从 U 型神经网络模型的瓶颈层中，提取全局特征；解码子单元用于利用 U 型神经网络模型的编码器模块，对第一图像特征进行解码，
35 得到语义分割结果。

本申请上述实施例中，融合单元包括：切分子单元、第三提取子单元和融合子单元。

其中，切分子单元用于分别对语义分割结果和图像进行切分，得到多个子分割结果和多个子图像；第三提取子单元用于分别对多个子分割结果和多个子图像进行特征提取，得到多个子分割结果的子分割特征，以及多个子图像的子图像特征；融合子单元用于将子分割特征和子图像特征进行融合，得到融合特征。

本申请上述实施例中，第一处理单元包括：拼接子单元和处理子单元。

其中，拼接子单元用于将全局特征和融合特征进行拼接，得到第一拼接特征；处理子单元用于利用自注意力模型对第一拼接特征进行自注意力处理，得到第一区域特征。

本申请上述实施例中，第一确定模块包括：第二处理单元。

其中，第二处理单元用于利用交叉注意力模型对第一区域特征和多个原型进行注意力处理，得到第二区域特征。

本申请上述实施例中，第一确定模块还包括：获取单元、聚类单元和构建单元。

其中，获取单元用于获取不同监测区域的全局特征；聚类单元用于对不同监测区域的全局特征进行聚类，得到多个特征集合；构建单元用于基于多个特征集合的中心特征，构建多个原型。

本申请上述实施例中，在基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征之后，该装置还包括：第三确定模块、第一更新模块和第二更新模块。

其中，第三确定模块用于从预先构建的多个原型中，确定与第二区域特征匹配成功的目标原型；第一更新模块用于对目标原型进行动量更新，得到更新后的区域特征；第二更新模块用于基于更新后的区域特征对多个原型进行更新。

本申请上述实施例中，第二确定模块包括：第三处理单元、第四处理单元、第五处理单元和汇总单元。

其中，第三处理单元用于基于全局特征识别监测区域的特征信息，得到第一子识别结果；第四处理单元用于基于第一区域特征识别监测区域的特征信息，得到第二子识别结果；第五处理单元用于基于第二区域特征识别监测区域的特征信息，得到第三子识别结果；汇总单元用于对第一子识别结果、第二子识别结果和第三子识别结果进行汇总，得到识别结果。

需要说明的是，本申请上述实施例中涉及到的优选实施方案与实施例1提供的方案以及应用场景、实施过程相同，但不仅限于实施例1所提供的方案。

实施例10

根据本申请实施例，还提供了一种用于实施上述图像处理方法的图像处理装置。

图12是根据本申请实施例7的一种图像处理装置的结构示意图，如图12所示，该装置包括：第一显示模块1202和第二显示模块1204。

其中，第一显示模块用于响应作用于操作界面上的输入指令，在操作界面上显示多张

图像，其中，图像的显示内容至少包含待监测对象的目标部位的监测区域；第二显示模块用于响应作用于操作界面上的图像处理指令，在操作界面上显示监测区域的识别结果，其中，识别结果是基于第一区域特征和第二区域特征识别监测区域的特征信息得到的，第二区域特征是基于预先构建的多个原型和第一区域特征之间的依赖关系确定的，第一区域特征是对医学图像进行语义分割得到的。

此处需要说明的是，上述第一显示模块和第二显示模块对应于实施例 2 中的步骤 S502 至步骤 S504，两个模块与对应的步骤所实现的实例和应用场景相同，但不限于上述实施例 2 所公开的内容。需要说明的是，上述模块或单元可以是存储在存储器中并由一个或多个处理器处理的硬件组件或软件组件，上述模块也可以作为装置的一部分可以运行在实施例 1 提供的 AR/VR 设备中。

需要说明的是，本申请上述实施例中涉及到的优选实施方案与实施例 1 提供的方案以及应用场景、实施过程相同，但不仅限于实施例 1 所提供的方案。

实施例 11

根据本申请实施例，还提供了一种用于实施上述图像处理方法的图像处理装置。图 13 是根据本申请实施例 8 的一种图像处理装置的结构示意图，如图 13 所示，该装置包括：显示模块 1302、分割模块 1304、第一确定模块 1306、第二确定模块 1308 和驱动模块 13010。

其中，显示模块用于在虚拟现实 VR 设备或增强现实 AR 设备的呈现画面上展示多张图像，其中，图像的显示内容至少包含待监测对象的目标部位的监测区域；分割模块用于对图像进行语义分割，得到图像中监测区域的第一区域特征；第一确定模块用于基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；第二确定模块用于基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果；驱动模块用于驱动 VR 设备或 AR 设备渲染展示识别结果。

此处需要说明的是，上述显示模块、分割模块、第一确定模块、第二确定模块和驱动模块对应于实施例 3 中的步骤 S702 至步骤 S7010，五个模块与对应的步骤所实现的实例和应用场景相同，但不限于上述实施例 3 所公开的内容。需要说明的是，上述模块或单元可以是存储在存储器中并由一个或多个处理器处理的硬件组件或软件组件，上述模块也可以作为装置的一部分可以运行在实施例 1 提供的 AR/VR 设备中。

需要说明的是，本申请上述实施例中涉及到的优选实施方案与实施例 1 提供的方案以及应用场景、实施过程相同，但不仅限于实施例 1 所提供的方案。

实施例 12

根据本申请实施例，还提供了一种用于实施上述图像处理方法的图像处理装置。图 14 是根据本申请实施例 9 的一种图像处理装置的结构示意图，如图 14 所示，该装置包括：获取模块 1402、分割模块 1404、第一确定模块 1406、第二确定模块 1408 和输出模块 14010。

其中，获取模块，用于通过调用第一接口获取多张图像，其中，第一接口包括第一参

数，第一参数的参数值为多张图像，图像的显示内容至少包含待监测对象的目标部位的监测区域；分割模块，用于对图像进行语义分割，得到图像中监测区域的第一区域特征；第一确定模块，用于基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；第二确定模块，用于基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果；输出模块，用于通过调用第二接口输出识别结果，其中，第二接口包括第二参数，第二参数的参数值为识别结果。

此处需要说明的是，上述获取模块、分割模块、第一确定模块、第二确定模块和输出模块对应于实施例 4 中的步骤 S802 至步骤 S8010，五个模块与对应的步骤所实现的实例和应用场景相同，但不限于上述实施例 4 所公开的内容。需要说明的是，上述模块或单元可以是存储在存储器中并由一个或多个处理器处理的硬件组件或软件组件，上述模块也可以作为装置的一部分可以运行在实施例 1 提供的 AR/VR 设备中。

需要说明的是，本申请上述实施例中涉及到的优选实施方案与实施例 1 提供的方案以及应用场景、实施过程相同，但不仅限于实施例 1 所提供的方案。

实施例 13

根据本申请实施例，还提供了一种用于实施上述肺结节诊断方法的肺结节诊断装置。图 15 是根据本申请实施例 10 的一种肺结节诊断装置的结构示意图，如图 15 所示，该装置包括：获取模块 1502、分割模块 1504、确定模块 1506、诊断模块 1508。

其中，获取模块用于获取多张医学图像，其中，多张医学图像包含肺结节；分割模块用于对医学图像进行语义分割，得到医学图像中肺结节的第一结节特征；确定模块用于基于预先构建的多个原型和第一结节特征之间的依赖关系，确定肺结节的第二结节特征，其中，不同原型用于表征不同类型的肺结节；诊断模块，用于基于第一结节特征第二结节特征对肺结节进行诊断，得到肺结节的诊断结果，其中，诊断结果用于表征肺结节是良性结节或恶性结节。

此处需要说明的是，上述获取模块、分割模块、确定模块、诊断模块对应于实施例 5 中的步骤 S902 至步骤 S908，五个模块与对应的步骤所实现的实例和应用场景相同，但不限于上述实施例 5 所公开的内容。需要说明的是，上述模块或单元可以是存储在存储器中并由一个或多个处理器处理的硬件组件或软件组件，上述模块也可以作为装置的一部分可以运行在实施例 1 提供的 AR/VR 设备中。

需要说明的是，本申请上述实施例中涉及到的优选实施方案与实施例 1 提供的方案以及应用场景、实施过程相同，但不仅限于实施例 1 所提供的方案。

实施例 14

本申请的实施例可以提供一种 AR/VR 设备，该 AR/VR 设备可以是 AR/VR 设备群中的任意一个 AR/VR 设备。可选地，在本实施例中，上述 AR/VR 设备也可以替换为移动终端等终端设备。

可选地，在本实施例中，上述 AR/VR 设备可以位于计算机网络的多个网络设备中的至少一个网络设备。

在本实施例中，上述 AR/VR 设备可以执行图像处理方法中以下步骤的程序代码：在虚拟现实 VR 设备或增强现实 AR 设备的呈现画面上展示多张图像，其中，图像的显示内容至少包含待监测对象的目标部位的监测区域；对图像进行语义分割，得到图像中监测区域的第一区域特征；基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；基于第一区域特征和

第二区域特征识别监测区域的特征信息，确定监测区域的识别结果；驱动 VR 设备或 AR 设备渲染展示识别结果。

可选地，图 16 是根据本申请实施例的一种计算机终端的结构框图。如图 16 所示，该计算机终端 A 可以包括：一个或多个（图中仅示出一个）处理器 1602、存储器 1604、存储控制器、以及外设接口，其中，外设接口与射频模块、音频模块和显示器连接。

其中，存储器可用于存储软件程序以及模块，如本申请实施例中的图像处理方法和装置对应的程序指令/模块，处理器通过运行存储在存储器内的软件程序以及模块，从而执行各种功能应用以及数据处理，即实现上述的图像处理方法。存储器可包括高速随机存储器，还可以包括非易失性存储器，如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中，存储器可进一步包括相对于处理器远程设置的存储器，这些远程存储器可以通过网络连接至终端 A。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

处理器可以通过传输装置调用存储器存储的信息及应用程序，以执行下述步骤：获取多张图像，其中，图像的显示内容至少包含待监测对象的目标部位的监测区域；对图像进行语义分割，得到图像中监测区域的第一区域特征；基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果。

可选的，上述处理器还可以执行如下步骤的程序代码：对图像进行语义分割，得到图像的语义分割结果和全局特征；对语义分割结果和图像进行特征融合，得到融合特征；对全局特征和融合特征进行注意力处理，得到第一区域特征。

可选的，上述处理器还可以执行如下步骤的程序代码：利用 U 型神经网络模型的编码器模块对图像进行特征提取，得到图像的第一图像特征；从 U 型神经网络模型的瓶颈层中，提取全局特征；利用 U 型神经网络模型的编码器模块，对第一图像特征进行解码，得到语义分割结果。

可选的，上述处理器还可以执行如下步骤的程序代码：分别对语义分割结果和图像进行切分，得到多个子分割结果和多个子图像；分别对多个子分割结果和多个子图像进行特征提取，得到多个子分割结果的子分割特征，以及多个子图像的子图像特征；将子分割特

征和子图像特征进行融合，得到融合特征。

可选的，上述处理器还可以执行如下步骤的程序代码：将全局特征和融合特征进行拼接，得到第一拼接特征；利用自注意力模型对第一拼接特征进行自注意力处理，得到第一区域特征。

- 5 可选的，上述处理器还可以执行如下步骤的程序代码：利用交叉注意力模型对第一区域特征和多个原型进行注意力处理，得到第二区域特征。

可选的，上述处理器还可以执行如下步骤的程序代码：获取不同监测区域的全局特征；对不同监测区域的全局特征进行聚类，得到多个特征集合；基于多个特征集合的中心特征，构建多个原型。

- 10 可选的，上述处理器还可以执行如下步骤的程序代码：从预先构建的多个原型中，确定与第二区域特征匹配成功的目标原型；对目标原型进行动量更新，得到更新后的区域特征；基于更新后的区域特征对多个原型进行更新。

- 可选的，上述处理器还可以执行如下步骤的程序代码：基于全局特征识别监测区域的特征信息，得到第一子识别结果；基于第一区域特征识别监测区域的特征信息，得到第二子识别结果；基于第二区域特征识别监测区域的特征信息，得到第三子识别结果；对第一子识别结果、第二子识别结果和第三子识别结果进行汇总，得到识别结果。
- 15

- 采用本申请实施例，提供了一种获取多张图像；对图像进行语义分割，得到图像中监测区域的第一区域特征；基于预先构建的多个原型和第一区域特征之间的依赖关系，确定监测区域的第二区域特征；基于第一区域特征和第二区域特征识别监测区域的特征信息，确定监测区域的识别结果的方式。容易注意到的是，本申请不仅对图像进行了语义分割，得到区域特征，还结合了原型与区域特征之间的依赖关系，确保最终的区域特征更符合目标部位本身的属性，特征提取准确度更高，进而达到了更准确的对待监测对象的目标部位进行识别的目的，从而实现了提高待监测对象的目标部位的识别准确率的技术效果，进而解决了相关技术中对待监测图像进行识别的识别准确率低的技术问题。
- 20

- 25 本领域普通技术人员可以理解，图所示的结构仅为示意，计算机终端也可以是智能手机（如 Android 手机、iOS 手机等）、平板电脑、掌上电脑以及移动互联网设备（Mobile Internet Devices, MID）、PAD 等终端设备。图 16 并不对上述电子装置的结构造成限定。例如，计算机终端 A 还可包括比图 16 中所示更多或者更少的组件（如网络接口、显示装置等），或者具有与图 16 所示不同的配置。

- 30 本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通过程序来指令终端设备相关的硬件来完成，该程序可以存储于一计算机可读存储介质中，存储介质可以包括：闪存盘、只读存储器（Read-Only Memory, ROM）、随机存取器（Random Access Memory, RAM）、磁盘或光盘等。

实施例 14

- 35 本申请的实施例还提供了一种计算机可读存储介质。可选地，在本实施例中，上述计

计算机可读存储介质可以用于保存上述实施例 1 所提供的图像处理方法所执行的程序代码。

可选地,在本实施例中,上述计算机可读存储介质可以位于 AR/VR 设备网络中 AR/VR 设备终端群中的任意一个计算机终端中,或者位于移动终端群中的任意一个移动终端中。

5 可选地,在本实施例中,计算机可读存储介质被设置为存储用于执行以下步骤的程序代码:获取多张图像,其中,图像的显示内容至少包含待监测对象的目标部位的监测区域;对图像进行语义分割,得到图像中监测区域的第一区域特征;基于预先构建的多个原型和第一区域特征之间的依赖关系,确定监测区域的第二区域特征,其中,不同原型用于表征不同类型的监测区域;基于第一区域特征和第二区域特征识别监测区域的特征信息,确定监测区域的识别结果。

10 可选地,在本实施例中,计算机可读存储介质还被设置为存储用于执行以下步骤的程序代码:对图像进行语义分割,得到图像的语义分割结果和全局特征;对语义分割结果和图像进行特征融合,得到融合特征;对全局特征和融合特征进行注意力处理,得到第一区域特征。

15 可选地,在本实施例中,计算机可读存储介质还被设置为存储用于执行以下步骤的程序代码:利用 U 型神经网络模型的编码器模块对图像进行特征提取,得到图像的第一图像特征;从 U 型神经网络模型的瓶颈层中,提取全局特征;利用 U 型神经网络模型的编码器模块,对第一图像特征进行解码,得到语义分割结果。

20 可选地,在本实施例中,计算机可读存储介质还被设置为存储用于执行以下步骤的程序代码:分别对语义分割结果和图像进行切分,得到多个子分割结果和多个子图像;分别对多个子分割结果和多个子图像进行特征提取,得到多个子分割结果的子分割特征,以及多个子图像的子图像特征;将子分割特征和子图像特征进行融合,得到融合特征。

可选地,在本实施例中,计算机可读存储介质还被设置为存储用于执行以下步骤的程序代码:将全局特征和融合特征进行拼接,得到第一拼接特征;利用自注意力模型对第一拼接特征进行自注意力处理,得到第一区域特征。

25 可选地,在本实施例中,计算机可读存储介质还被设置为存储用于执行以下步骤的程序代码:利用交叉注意力模型对第一区域特征和多个原型进行注意力处理,得到第二区域特征。

30 可选地,在本实施例中,计算机可读存储介质还被设置为存储用于执行以下步骤的程序代码:获取不同监测区域的全局特征;对不同监测区域的全局特征进行聚类,得到多个特征集合;基于多个特征集合的中心特征,构建多个原型。

可选地,在本实施例中,计算机可读存储介质还被设置为存储用于执行以下步骤的程序代码:从预先构建的多个原型中,确定与第二区域特征匹配成功的目标原型;对目标原型进行动量更新,得到更新后的区域特征;基于更新后的区域特征对多个原型进行更新。

35 可选地,在本实施例中,计算机可读存储介质还被设置为存储用于执行以下步骤的程序代码:基于全局特征识别监测区域的特征信息,得到第一子识别结果;基于第一区域特

征识别监测区域的特征信息，得到第二子识别结果；基于第二区域特征识别监测区域的特征信息，得到第三子识别结果；对第一子识别结果、第二子识别结果和第三子识别结果进行汇总，得到识别结果。

实施例 15

5 本申请的实施例还提供了一种计算机程序产品，包括计算机程序，所述计算机程序在计算机中执行时，令计算机执行本申请的实施例提供的方法。

实施例 16

10 本申请的实施例还提供了一种计算机程序，其中，当所述计算机程序在计算机中执行时，令计算机执行本申请的实施例提供的方法。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。

在本申请的上述实施例中，对各个实施例的描述都各有侧重，某个实施例中未详述的部分，可以参见其他实施例的相关描述。

15 在本申请所提供的几个实施例中，应该理解到，所揭露的技术内容，可通过其它的方式实现。其中，以上所描述的装置实施例仅仅是示意性的，例如所述单元的划分，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，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，单元或模块的间接耦合或通信连接，可以是电性或其它的形式。

20 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元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单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元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，也可以采用软件功能单元的形式实现。

25 所述集成的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。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或者该技术方案的全部或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中，包括若干指令用以使得一台计算机设备（可为个人计算机、服务器或者网络设备等）执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U 盘、只读存储器（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

30

35 以上所述仅是本申请的优选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本申请原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本申请的保护范围。

权 利 要 求 书

1. 一种图像处理方法，其特征在于，包括：

获取多张图像，其中，所述图像的显示内容至少包含待监测对象的目标部位的监测区域；

5 对所述图像进行语义分割，得到所述图像中所述监测区域的第一区域特征；

基于预先构建的多个原型和所述第一区域特征之间的依赖关系，确定所述监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；

基于所述第一区域特征和所述第二区域特征识别所述监测区域的特征信息，确定所述监测区域的识别结果。

10 2. 根据权利要求1所述的方法，其特征在于，对所述图像进行语义分割，得到所述图像中所述监测区域的第一区域特征，包括：

对所述图像进行语义分割，得到所述图像的语义分割结果和全局特征；

对所述语义分割结果和所述图像进行特征融合，得到融合特征；

对所述全局特征和所述融合特征进行注意力处理，得到所述第一区域特征。

15 3. 根据权利要求2所述的方法，其特征在于，对所述图像进行语义分割，得到所述图像的语义分割结果和全局特征，包括：

利用U型神经网络模型的编码器模块对所述图像进行特征提取，得到所述图像的第一图像特征；

从所述U型神经网络模型的瓶颈层中，提取所述全局特征；

20 利用所述U型神经网络模型的编码器模块，对所述第一图像特征进行解码，得到所述语义分割结果。

4. 根据权利要求2或3所述的方法，其特征在于，对所述语义分割结果和所述图像进行特征融合，得到融合特征，包括：

25 分别对所述语义分割结果和所述图像进行切分，得到多个子分割结果和多个子图像；

分别对所述多个子分割结果和所述多个子图像进行特征提取，得到所述多个子分割结果的子分割特征，以及所述多个子图像的子图像特征；

将所述子分割特征和所述子图像特征进行融合，得到所述融合特征。

30 5. 根据权利要求2至4任意一项所述的方法，其特征在于，对所述全局特征和所述融合特征进行注意力处理，得到所述第一区域特征，包括：

将所述全局特征和所述融合特征进行拼接，得到第一拼接特征；

利用自注意力模型对所述第一拼接特征进行自注意力处理，得到所述第一区域特征。

35 6. 根据权利要求1至5任意一项所述的方法，其特征在于，基于预先构建的多个原型和所述第一区域特征之间的依赖关系，确定所述监测区域的第二区域特征，包括：

利用交叉注意力模型对所述第一区域特征和所述多个原型进行注意力处理，得到所述第二区域特征。

7. 根据权利要求6所述的方法，其特征在于，所述方法还包括：

获取所述不同监测区域的全局特征；

5 对所述不同监测区域的全局特征进行聚类，得到多个特征集合；

基于所述多个特征集合的中心特征，构建所述多个原型。

8. 根据权利要求1至7任意一项所述的方法，其特征在于，在基于预先构建的多个原型和所述第一区域特征之间的依赖关系，确定所述监测区域的第二区域特征之后，所述方法还包括：

10 从预先构建的多个原型中，确定与第二区域特征匹配成功的目标原型；

对所述目标原型进行动量更新，得到更新后的区域特征；

基于所述更新后的区域特征对所述多个原型进行更新。

9. 根据权利要求1至8任意一项所述的方法，其特征在于，基于所述第一区域特征和所述第二区域特征识别所述监测区域的特征信息，确定所述监测区域的识别结果，包

15 括：

基于全局特征识别所述监测区域的特征信息，得到第一子识别结果；

基于所述第一区域特征识别所述监测区域的特征信息，得到第二子识别结果；

基于所述第二区域特征识别所述监测区域的特征信息，得到第三子识别结果；

对所述第一子识别结果、所述第二子识别结果和所述第三子识别结果进行汇总，得

20 到所述识别结果。

10. 一种图像处理方法，其特征在于，包括：

响应作用于操作界面上的输入指令，在所述操作界面上显示多张图像，其中，所述图像的显示内容至少包含待监测对象的目标部位的监测区域；

25 响应作用于所述操作界面上的图像处理指令，在所述操作界面上显示所述监测区域的识别结果，其中，所述识别结果是基于第一区域特征和第二区域特征识别所述监测区域的特征信息得到的，所述第二区域特征是基于预先构建的多个原型和所述第一区域特征之间的依赖关系确定的，所述第一区域特征是对医学图像进行语义分割得到的。

11. 一种图像处理方法，其特征在于，包括：

30 通过调用第一接口获取多张图像，其中，所述第一接口包括第一参数，所述第一参数的参数值为所述多张图像，所述图像的显示内容至少包含待监测对象的目标部位的监测区域；

对所述图像进行语义分割，得到所述图像中所述监测区域的第一区域特征；

基于预先构建的多个原型和所述第一区域特征之间的依赖关系，确定所述监测区域的第二区域特征，其中，不同原型用于表征不同类型的监测区域；

35 基于所述第一区域特征和所述第二区域特征识别所述监测区域的特征信息，确定所

述监测区域的识别结果;

通过调用第二接口输出所述识别结果, 其中, 所述第二接口包括第二参数, 所述第二参数的参数值为所述识别结果。

12. 一种癌症计算机辅助诊断方法, 包括:

5 获取多张医学图像, 其中, 所述医学图像包含待监测对象的目标监测部位;

对所述医学图像进行语义分割, 得到所述医学图像中所述监测部位的第一部位特征;

基于预先构建的多个原型和所述第一部位特征之间的依赖关系, 确定所述监测部位的第二部位特征, 其中, 不同原型用于表征不同类型的监测部位;

10 基于所述第一部位特征和所述第二部位特征对所述监测部位进行诊断, 得到所述监测部位的诊断结果, 其中, 所述诊断结果用于表征所述监测部位是恶性病症或良性病症。

13. 一种肺癌计算机辅助诊断方法, 包括:

获取多张医学图像, 其中, 所述医学图像包含肺部病变;

15 对所述医学图像进行语义分割, 得到所述医学图像中肺部病变的第一病变特征;

基于预先构建的多个原型和所述第一结节特征之间的依赖关系, 确定所述肺部病变的第二病变特征, 其中, 不同原型用于表征不同类型的肺部病变;

基于所述第一病变特征和所述第二病变特征对所述肺结节进行诊断, 得到所述肺部病变的诊断结果, 其中, 所述诊断结果用于表征所述肺部病变是良性病症或恶性病症。

20 14. 一种图像处理装置, 包括:

获取模块, 用于获取多张图像, 其中, 所述图像的显示内容至少包含待监测对象的目标部位的监测区域;

分割模块, 用于对所述图像进行语义分割, 得到所述图像中所述监测区域的第一区域特征;

25 第一确定模块, 用于基于预先构建的多个原型和所述第一区域特征之间的依赖关系, 确定所述监测区域的第二区域特征, 其中, 不同原型用于表征不同类型的监测区域;

第二确定模块, 基于所述第一区域特征和所述第二区域特征识别所述监测区域的特征信息, 确定所述监测区域的识别结果。

30 15. 一种图像处理装置, 包括:

第一显示模块, 用于响应作用于操作界面上的输入指令, 在所述操作界面上显示多张图像, 其中, 所述图像的显示内容至少包含待监测对象的目标部位的监测区域;

第二显示模块, 用于响应作用于所述操作界面上的图像处理指令, 在所述操作界面上显示所述监测区域的识别结果, 其中, 所述识别结果是基于第一区域特征和第二区域特征识别所述监测区域的特征信息得到的, 所述第二区域特征是基于预先构建的多个原

35

型和所述第一区域特征之间的依赖关系确定的，所述第一区域特征是对医学图像进行语义分割得到的。

- 5 16. 一种癌症的计算机辅助诊断系统，包括存储器、处理器以及存储在所述存储器上并在所述处理器上运行的计算机程序，所述处理器执行所述计算机程序可用于执行一种癌症计算机辅助诊断方法，所述方法包括：

获取多张医学图像，其中，所述医学图像包含待监测对象的目标监测部位；

对所述医学图像进行语义分割，得到所述医学图像中所述监测部位的第一部位特征；

- 10 基于预先构建的多个原型和所述第一部位特征之间的依赖关系，确定所述监测部位的第二部位特征，其中，不同原型用于表征不同类型的监测部位；

基于所述第一部位特征和所述第二部位特征对所述监测部位进行诊断，得到所述监测部位的诊断结果，其中，所述诊断结果用于表征所述监测部位是恶性病症或良性病症。

- 15 17. 一种电子设备，其特征在于，包括：

存储器，存储有可执行程序；

处理器，用于运行所述程序，其中，所述程序运行时执行权利要求 1 至 13 中任意一项所述的方法。

- 20 18. 一种计算机可读存储介质，其特征在于，所述计算机可读存储介质包括存储的可执行程序，其中，在所述可执行程序运行时控制所述计算机可读存储介质所在设备执行权利要求 1 至 13 中任意一项所述的方法。

19. 一种计算机程序产品，包括计算机可执行指令，该计算机可执行指令被处理器执行时实现权利要求 1 至 13 任意一项所述方法的步骤。

20. 一种计算机程序，该计算机程序被处理器执行时实现权利要求 1 至 13 任意一项所述方法的步骤。

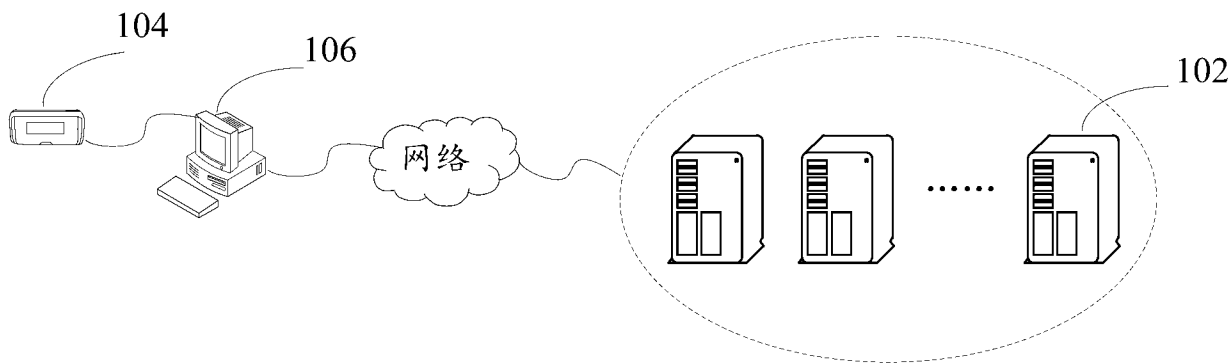


图 1

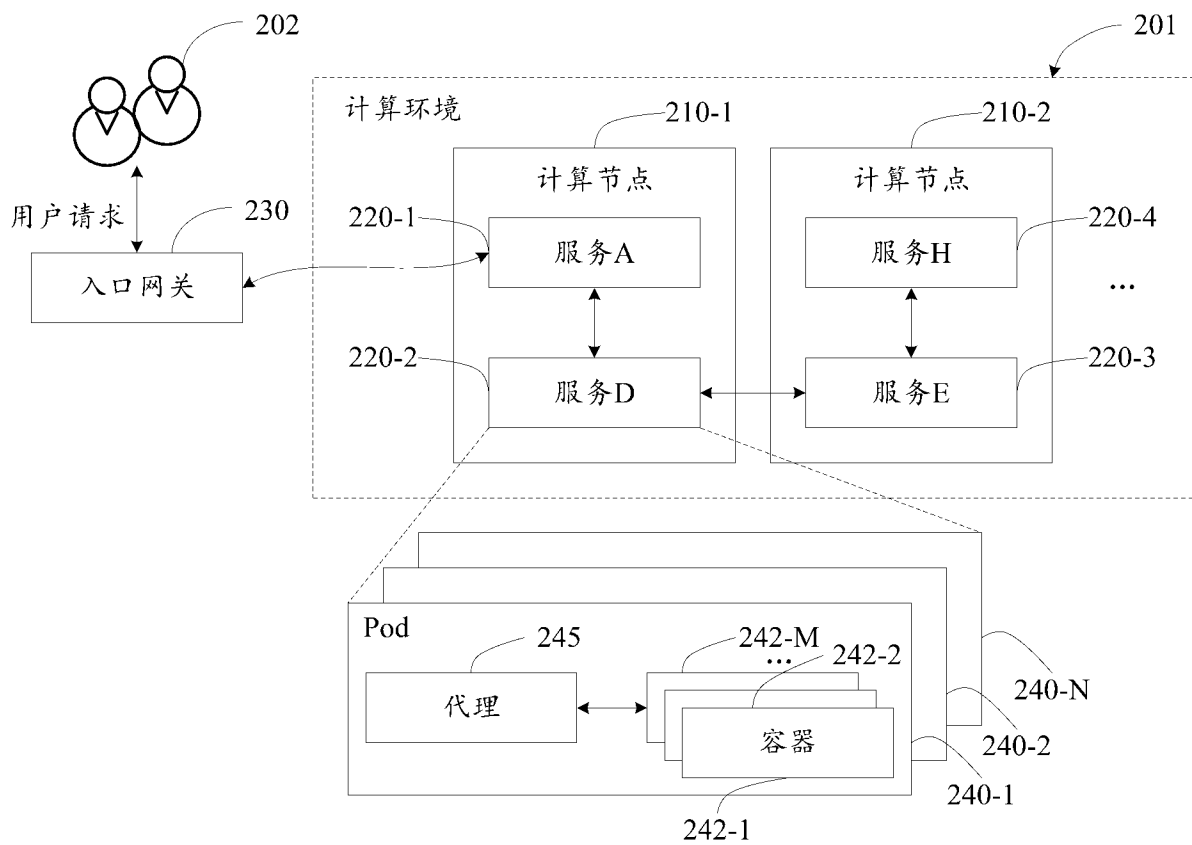


图 2

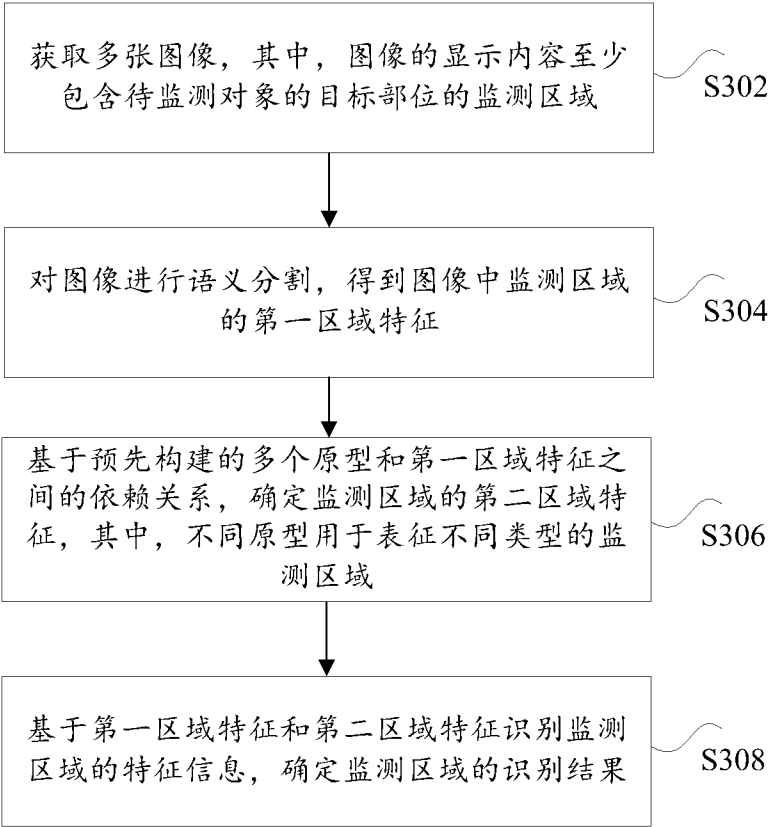


图 3

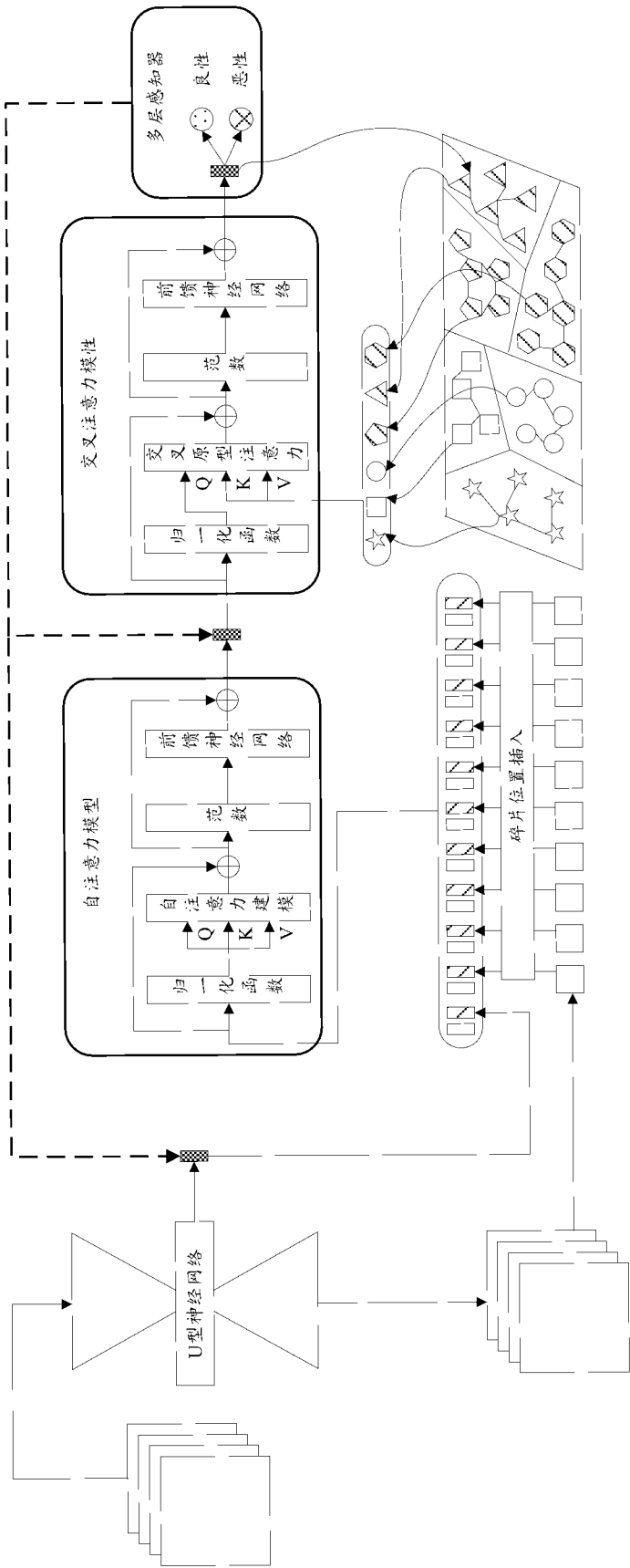


图 4

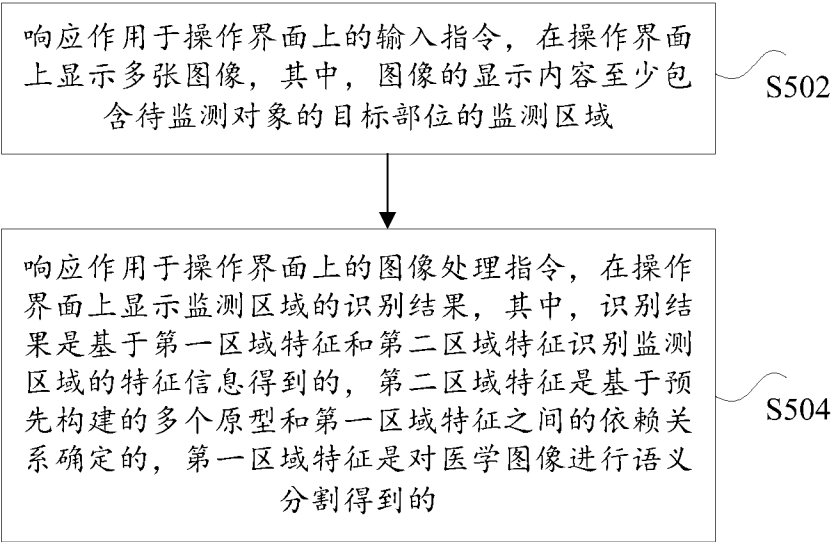


图 5

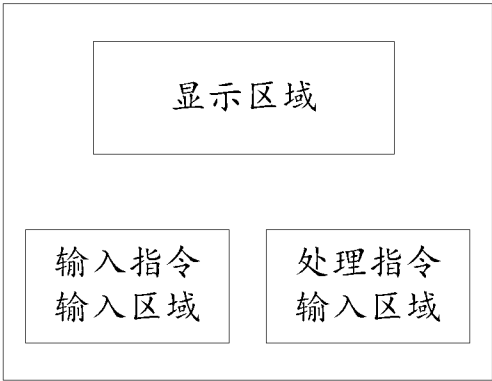


图 6

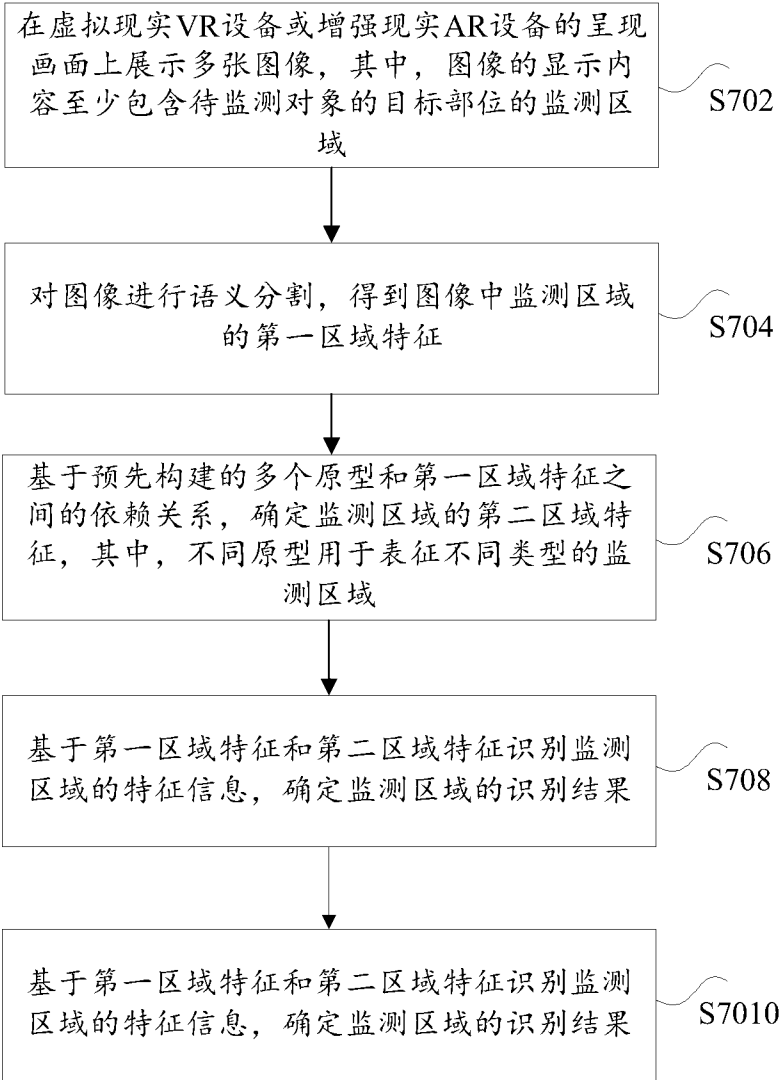


图 7

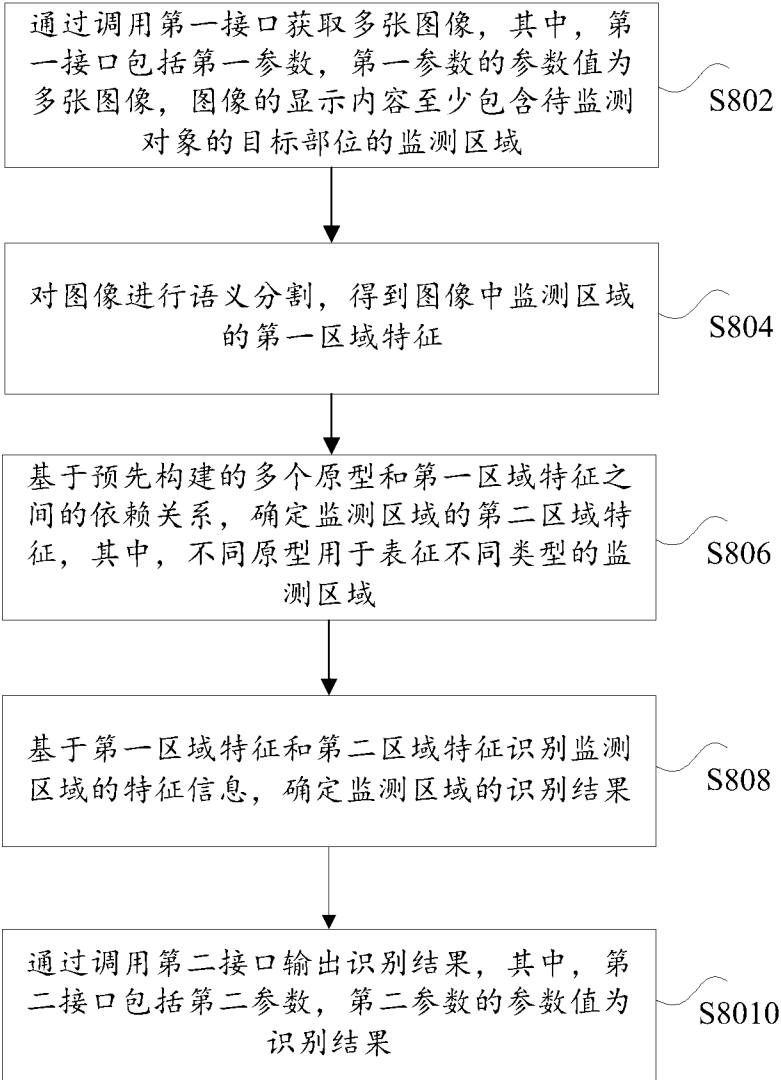


图 8

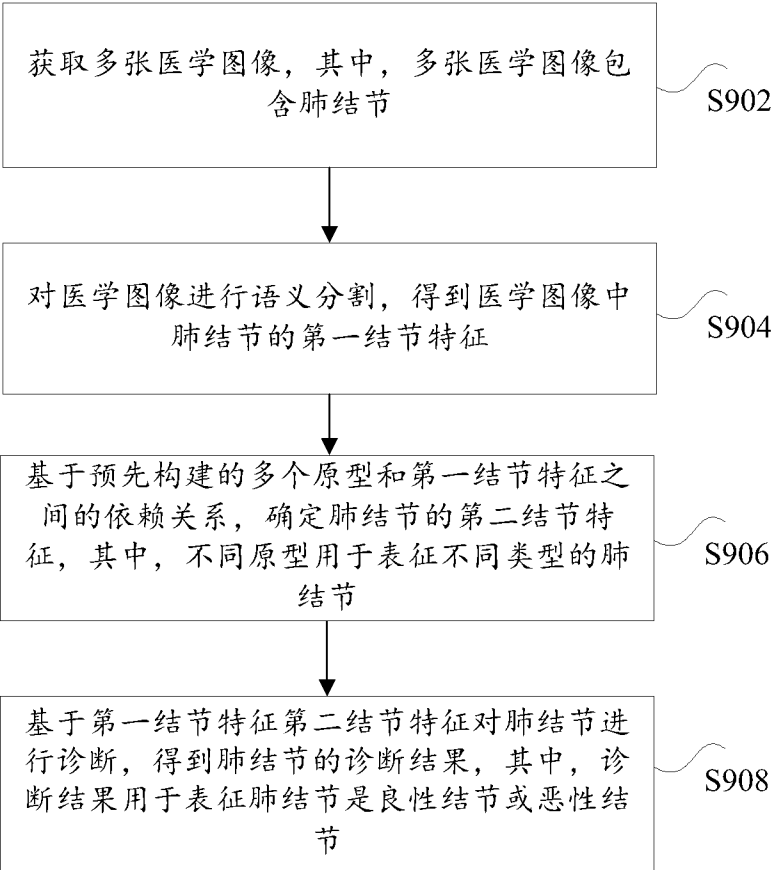


图 9

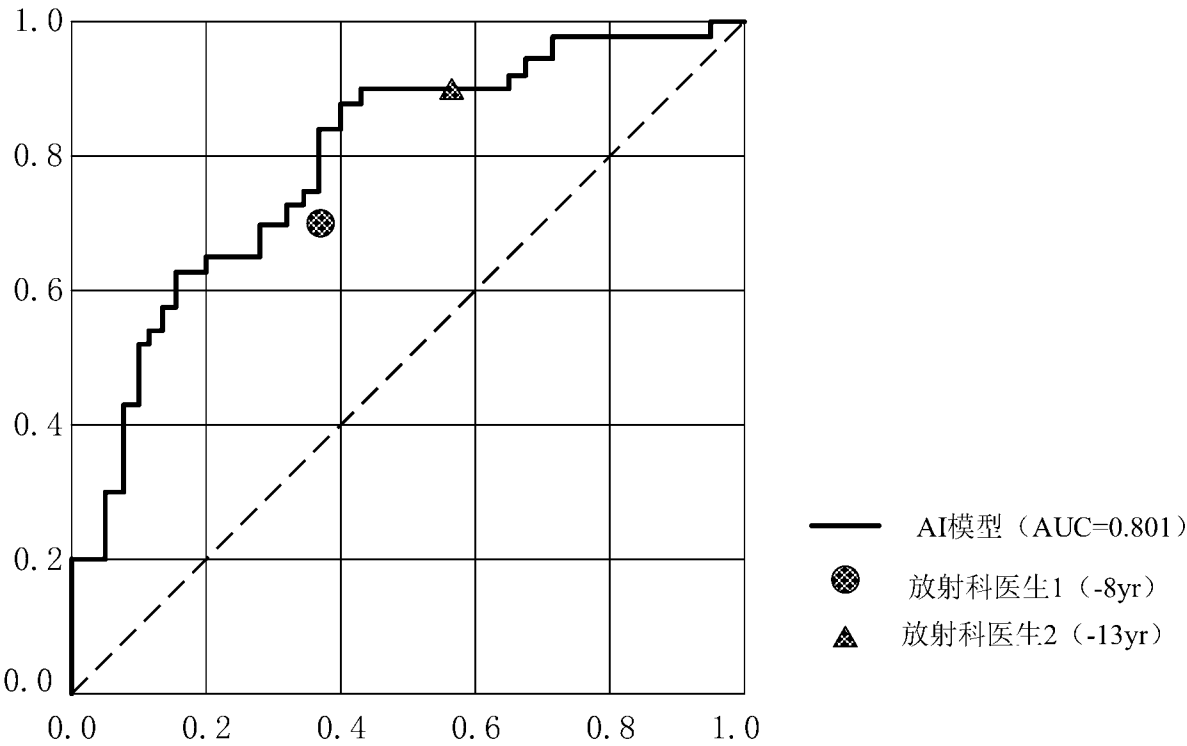


图 10



图 11

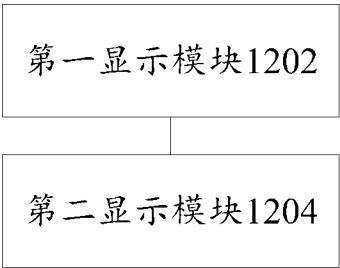


图 12



图 13



图 14



图 15

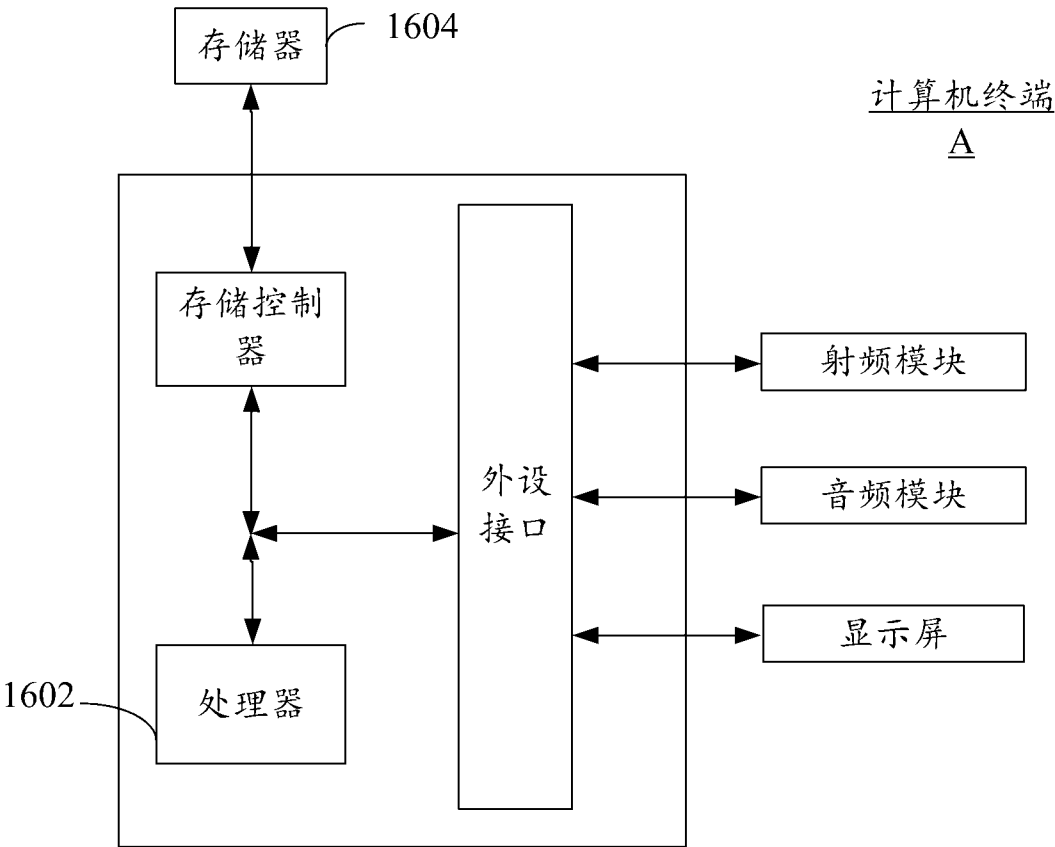


图 16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/103723

A. CLASSIFICATION OF SUBJECT MATTER G06V 20/20(2022.01)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) IPC:G06V,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) VEN, CNABS, CNTXT, WOTXT, EPTXT, USTXT, CNKI, IEEE: 图像, 图象, 识别, 检测, 监控, 区域, 感兴趣, 特征, 语义, 分割, 依赖, 关系, 其他, 交叉, 注意, 关注, 肺, 结节, 癌, 医学, 医疗, images, identify, detect, region, ROI, feature, semantic, segment, depend, relation, attention, cross, pulmonary, nodule, cancer, medical		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 117095320 A (ALIBABA DAMOYUAN (HANGZHOU) TECHNOLOGY CO., LTD.) 21 November 2023 (2023-11-21) claims 1-14, and description, paragraphs [0134]-[0171]	1-20
A	CN 116188392 A (ALIBABA (CHINA) CO., LTD.) 30 May 2023 (2023-05-30) description, paragraphs [0004]-[0253], and figures 1-13	1-20
A	WO 2022248727 A1 (DEEPMIND TECHNOLOGIES LTD.) 01 December 2022 (2022-12-01) entire document	1-20
A	US 10430946 B1 (INCEPTION INSTITUTE OF ARTIFICIAL INTELLIGENCE, LTD.) 01 October 2019 (2019-10-01) entire document	1-20
A	CN 114581459 A (ZHEJIANG UNIVERSITY) 03 June 2022 (2022-06-03) entire document	1-20
A	CN 114612902 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 10 June 2022 (2022-06-10) 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 “D” document cited by the applicant in the international application “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 13 September 2024		Date of mailing of the international search report 20 September 2024
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/CN) China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/103723

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	117095320	A	21 November 2023	None			
CN	116188392	A	30 May 2023	WO	2024141092	A1	04 July 2024
WO	2022248727	A1	01 December 2022	AU	2022281121	A1	12 October 2023
				US	2024232580	A1	11 July 2024
				JP	2024521621	A	04 June 2024
				CA	3217060	A1	01 December 2022
				KR	20230159585	A	21 November 2023
				EP	4302237	A1	10 January 2024
				CN	117242454	A	15 December 2023
				IN	202327067582	A	29 March 2024
US	10430946	B1	01 October 2019	WO	2020183230	A1	17 September 2020
CN	114581459	A	03 June 2022	None			
CN	114612902	A	10 June 2022	HK	40070352	A0	28 October 2022

A. 主题的分类

G06V 20/20(2022.01)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)

IPC:G06V,G06F

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))

VEN, CNABS, CNTXT, WOTXT, EPTXT, USTXT, CNKI, IEEE: 图像, 图象, 识别, 检测, 监控, 区域, 感兴趣, 特征, 语义, 分割, 依赖, 关系, 其他, 交叉, 注意, 关注, 肺, 结节, 癌, 医学, 医疗, images, identify, detect, region, ROI, feature, semantic, segment, depend, relation, attention, cross, pulmonary, nodule, cancer, medical

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
PX	CN 117095320 A (阿里巴巴达摩院(杭州)科技有限公司) 2023年11月21日 (2023 - 11 - 21) 权利要求1-14, 说明书第[0134]-[0171]段	1-20
A	CN 116188392 A (阿里巴巴(中国)有限公司) 2023年5月30日 (2023 - 05 - 30) 说明书第[0004]-[0253]段、图1-13	1-20
A	WO 2022248727 A1 (DEEPMIND TECHNOLOGIES LIMITED) 2022年12月1日 (2022 - 12 - 01) 全文	1-20
A	US 10430946 B1 (INCEPTION INSTITUTE OF ARTIFICAL INTELLIGENCE, LTD.) 2019年10月1日 (2019 - 10 - 01) 全文	1-20
A	CN 114581459 A (浙江大学) 2022年6月3日 (2022 - 06 - 03) 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

“D” 申请人在国际申请中引证的文件

“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

2024年9月13日

国际检索报告邮寄日期

2024年9月20日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

董洪梅

电话号码 (+86) 010-53961528

PCT/ISA/210 表(第2页) (2022年7月)

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
A	CN 114612902 A (腾讯科技（深圳）有限公司) 2022年6月10日 (2022 - 06 - 10) 全文	1-20

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/103723

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	117095320	A	2023年11月21日	无			
CN	116188392	A	2023年5月30日	WO	2024141092	A1	2024年7月4日
WO	2022248727	A1	2022年12月1日	AU	2022281121	A1	2023年10月12日
				US	2024232580	A1	2024年7月11日
				JP	2024521621	A	2024年6月4日
				CA	3217060	A1	2022年12月1日
				KR	20230159585	A	2023年11月21日
				EP	4302237	A1	2024年1月10日
				CN	117242454	A	2023年12月15日
				IN	202327067582	A	2024年3月29日
US	10430946	B1	2019年10月1日	WO	2020183230	A1	2020年9月17日
CN	114581459	A	2022年6月3日	无			
CN	114612902	A	2022年6月10日	HK	40070352	A0	2022年10月28日