



(12) 发明专利

(10) 授权公告号 CN 113689954 B

(45) 授权公告日 2024.10.18

(21) 申请号 202110977285.4

(22) 申请日 2021.08.24

(65) 同一申请的已公布的文献号
申请公布号 CN 113689954 A

(43) 申请公布日 2021.11.23

(73) 专利权人 平安科技(深圳)有限公司
地址 518000 广东省深圳市福田区福田街
道福安社区益田路5033号平安金融中
心23楼

(72) 发明人 柳杨 张潇月 吕彬 吕传峰

(74) 专利代理机构 上海汉之律师事务所 31378
专利代理师 冯华

(51) Int.Cl.
G16H 50/30 (2018.01)
G16H 50/70 (2018.01)
G16H 50/20 (2018.01)
G16H 30/40 (2018.01)

G16H 10/60 (2018.01)
G06V 10/764 (2022.01)
G06V 10/82 (2022.01)
G06V 10/44 (2022.01)
G06V 10/52 (2022.01)
G06T 7/00 (2017.01)
G06T 7/62 (2017.01)
G06T 5/90 (2024.01)
G06N 3/0464 (2023.01)
G06N 3/0455 (2023.01)
A61B 5/00 (2006.01)
A61B 3/14 (2006.01)
A61B 3/12 (2006.01)

(56) 对比文件

CN 111048210 A, 2020.04.21
CN 111340789 A, 2020.06.26
CN 112716446 A, 2021.04.30

审查员 刘自琴

权利要求书3页 说明书15页 附图5页

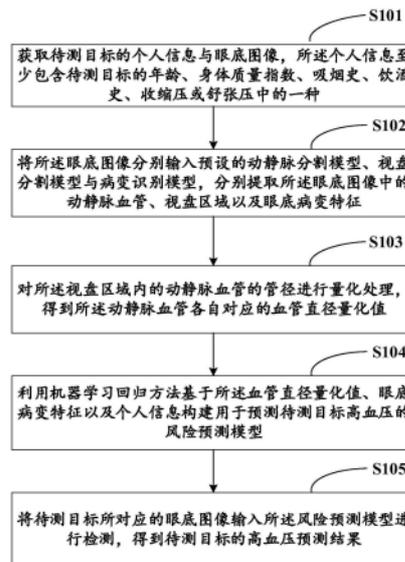
(54) 发明名称

高血压风险预测方法、装置、设备及介质

(57) 摘要

本发明涉及医疗科技领域,提出一种高血压风险预测方法、装置、设备及介质,该方法通过获取待测目标的个人信息与眼底图像;将眼底图像分别输入预设的动静脉分割模型、视盘分割模型与病变识别模型,分别提取眼底图像中的动静脉血管、视盘区域以及眼底病变特征;对视盘区域内的动静脉血管的管径进行量化处理,得到动静脉血管各自对应的血管直径量化值;利用机器学习回归方法基于血管直径量化值、眼底病变特征以及个人信息构建用于预测待测目标高血压的风险预测模型;将待测目标所对应的眼底图像输入风险预测模型进行检测,得到待测目标的高血压预测结果。由于风险预测模型采用多个变量因子进行训练,极大提高了高血压风险预测的准确性。

CN 113689954 B



1. 一种高血压风险预测方法,其特征在于,所述方法包括:

获取待测目标的个人信息与眼底图像,所述个人信息至少包含待测目标的年龄、身体质量指数、吸烟史、饮酒史、收缩压或舒张压中的一种;

将所述眼底图像分别输入预设的动静脉分割模型、视盘分割模型与病变识别模型,分别提取所述眼底图像中的动静脉血管、视盘区域以及眼底病变特征,其中,采用UNET网络基于训练集分别构建动静脉分割模型和视盘分割模型,所述病变识别模型根据不同的病变症状采用不同的神经网络模型进行训练,其中,所述病变症状至少包括动静脉压迫征、视乳头水肿、出血和渗出;其中,所述UNET网络包括编码层与解码层,所述编码层由卷积层与下采样层构成,用于提取特征;所述解码层由卷积层与下采样层构成,用于识别提取的特征;用于提取特征的UNET网络的损失函数为Dice Loss;用于识别特征的UNET网络的损失函数为多类交叉熵损失;

$$L_s = 1 - \frac{2|x \cap z| + 1}{|x| + |z| + 1} \quad \text{式 (1)}$$

$$L(y, y', y_1, y_1') = -(y \ln y' + (1-y) \ln(1-y')) - (y_1 \ln y_1' + (1-y_1) \ln(1-y_1')) \quad \text{式 (2)}$$

式(1)中, L_s 为损失函数Dice Loss, $|x \cap z|$ 表示眼底图像 x 和 z 之间的交集, $|x|$ 和 $|z|$ 分别表示 x 和 z 的元素的个数;式(2)中,多类交叉熵损失为 $L(y, y', y_1, y_1')$, y 表示眼底图像 x 真实动静脉标注概率图的概率, y_1 表示眼底图像 x 动静脉边界的真实标注概率图的概率, y' 表示眼底图像 x 真实动静脉标注概率图经 $\text{Loc}(\cdot)$ 的输出, y_1' 表示眼底图像 x 动静脉边界的真实标注概率图经 $\text{Loc}(\cdot)$ 的输出;

对所述视盘区域内的动静脉血管的管径进行量化处理,得到所述动静脉血管各自对应的血管直径量化值;

利用机器学习回归方法基于所述血管直径量化值、眼底病变特征以及个人信息构建用于预测待测目标高血压的风险预测模型;其中,利用回归分析逐一分析所述血管直径量化值、眼底病变特征以及个人信息,得到相关性满足预设值的变量因子;根据各个变量因子与高血压的相关性,按照主成分分析法确定各个变量因子的权重系数;将所述变量因子以及对应的权重系数输入基于机器学习的多因素回归模型进行训练,得到预测待测目标高血压的风险预测模型;

$$\text{Risk}_{\text{高血压}} = \frac{1}{1 + e^{-(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}} \quad \text{式 (3)}$$

式(3)中, x_1, \dots, x_n 为上述变量因子, β_1, \dots, β_n 为各变量因子的权重,该风险值的大小表明各变量因子每增加1个单位,其患病风险增至 e^β 倍;

将待测目标所对应的眼底图像输入所述风险预测模型进行检测,得到待测目标的高血压预测结果。

2. 如权利要求1所述的高血压风险预测方法,其特征在于,所述采用UNET网络基于训练集分别构建动静脉分割模型和视盘分割模型包括:

基于训练集采用两个UNET网络训练得到动静脉分割模型和视盘分割模型;所述动静脉分割模型中一个UNET网络用于提取所述眼底图像内的血管组织特征,另一个UNET网络用于根据所述血管组织特征输出概率确定所述血管组织为动脉或静脉;所述视盘分割模型中一

个UNET网络用于提取所述眼底图像内的视盘组织特征,另一个UNET网络用于根据所述视盘组织特征输出概率确定所述视盘组织是否为视盘。

3.如权利要求1所述的高血压风险预测方法,其特征在于,所述病变识别模型根据不同的病变症状采用不同的神经网络模型进行训练的步骤,还包括:

将所述待测目标的眼底图像按照病变症状不同分别进行标注,分别构成训练不同神经网络模型的多个数据集,多个所述数据集至少包括第一数据集、第二数据集、第三数据集与第四数据集;

基于所述第一数据集内标注的动静脉压迫征病变区域和非动静脉压迫征病变区域,采用所述第一数据集训练卷积神经网络得到动静脉压迫征识别模型;

基于所述第二数据集内标注的动静脉压迫征病变区域和非动静脉压迫征病变区域,采用所述第二数据集训练卷积神经网络得到视乳头水肿识别模型;

基于所述第三数据集内标注的出血病变区域和非出血病变区域第三训练数据,采用所述第三数据集训练全卷积深度学习网络得到出血病变识别模型;

基于所述第四数据集内标注的渗出病变区域和非渗出病变区域,采用所述第四数据集训练全卷积深度学习网络得到动静脉压迫征识别模型。

4.如权利要求1所述的高血压风险预测方法,其特征在于,所述将所述眼底图像分别输入预设的动静脉分割模型、视盘分割模型与病变识别模型之前,包括:

将眼底图像转换成预设规格大小的图像;

对预设规格的所述眼底图像进行归一化处理,得到像素色彩均匀的眼底图像;

对所述像素色彩均匀的眼底图像分别进行亮度、对比度、色彩的增强处理,得到增强后的眼底图像。

5.如权利要求1所述的高血压风险预测方法,其特征在于,所述对所述视盘区域内的动静脉血管的管径进行量化处理,得到所述动静脉血管各自对应的血管直径量化值还包括:

提取预设视盘直径范围内的动脉血管与静脉血管,采用修正的Parr-Hubbard公式分别计算视网膜中央动脉直径等效值和视网膜中央静脉直径等效值,得到动脉血管与静脉血管各自对应的血管直径量化值;

$$CRAE = \sqrt{0.88 \times (A_i^2 + A_j^2)}$$

$$CRVE = \sqrt{0.95 \times (V_i^2 + V_j^2)}$$

其中, A_i 和 A_j 分别为迭代获取到的最大动脉血管直径和最小动脉血管直径; V_i 和 V_j 分别为迭代获取到的最大静脉血管直径和最小静脉血管直径,CRAE和CRVE分别为视网膜中央动脉直径等效值和视网膜中央静脉直径等效值。

6.一种高血压风险预测装置,其特征在于,所述装置包括:

获取模块,用于获取待测目标的个人信息与眼底图像,所述个人信息至少包含待测目标的年龄、身体质量指数、吸烟史、饮酒史、收缩压或舒张压中的一种;

特征提取模块,用于将所述眼底图像分别输入预设的动静脉分割模型、视盘分割模型与病变识别模型,分别提取所述眼底图像中的动静脉血管、视盘区域以及眼底病变特征;其中,采用UNET网络基于训练集分别构建动静脉分割模型和视盘分割模型,所述病变识别模

型根据不同的病变症状采用不同的神经网络模型进行训练,其中,所述病变症状至少包括动静脉压迫征、视乳头水肿、出血和渗出;其中,所述UNET网络包括编码层与解码层,所述编码层由卷积层与下采样层构成,用于提取特征;所述解码层由卷积层与下采样层构成,用于识别提取的特征;用于提取特征的UNET网络的损失函数为Dice Loss;用于识别特征的UNET网络的损失函数为多类交叉熵损失;

$$L_s = 1 - \frac{2|x \cap z| + 1}{|x| + |z| + 1} \quad \text{式 (1)}$$

$$L(y, y', y_1, y_1') = -(y \ln y' + (1-y) \ln(1-y')) - (y_1 \ln y_1' + (1-y_1) \ln(1-y_1')) \quad \text{式 (2)}$$

式(1)中, L_s 为损失函数Dice Loss, $|x \cap z|$ 表示眼底图像 x 和 z 之间的交集, $|x|$ 和 $|z|$ 分别表示 x 和 z 的元素个数;式(2)中,多类交叉熵损失为 $L(y, y', y_1, y_1')$, y 表示眼底图像 x 真实动静脉标注概率图的概率, y_1 表示眼底图像 x 动静脉边界的真实标注概率图的概率, y' 表示眼底图像 x 真实动静脉标注概率图经 $\text{Loc}(\cdot)$ 的输出, y_1' 表示眼底图像 x 动静脉边界的真实标注概率图经 $\text{Loc}(\cdot)$ 的输出;

血管直径量化模块,用于对所述视盘区域内的动静脉血管的管径进行量化处理,得到所述动静脉血管各自对应的血管直径量化值;

风险预测模型构建模块,利用机器学习回归方法基于所述血管直径量化值、眼底病变特征以及个人信息构建用于预测待测目标高血压的风险预测模型;其中,利用回归分析逐一分析所述血管直径量化值、眼底病变特征以及个人信息,得到相关性满足预设值的变量因子;根据各个变量因子与高血压的相关性,按照主成分分析法确定各个变量因子的权重系数;将所述变量因子以及对应的权重系数输入基于机器学习的多因素回归模型进行训练,得到预测待测目标高血压的风险预测模型;

$$\text{Risk}_{\text{高血压}} = \frac{1}{1 + e^{-(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}} \quad \text{式 (3)}$$

式(3)中, x_1, \dots, x_n 为上述变量因子, β_1, \dots, β_n 为各变量因子的权重,该风险值的大小表明各变量因子每增加1个单位,其患病风险增至 e^β 倍;

风险预测模块,用于将待测目标所对应的眼底图像输入所述风险预测模型进行检测,得到待测目标的高血压预测结果。

7. 一种计算机设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序时实现权利要求1至5任一项所述的方法的步骤。

8. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现权利要求1至5中任一项所述的方法的步骤。

高血压风险预测方法、装置、设备及介质

技术领域

[0001] 本发明涉及医疗科技领域,提供一种高血压风险预测方法、装置、设备及介质。

背景技术

[0002] 随着我国经济的快速发展,生活水平在不断提高,饮食结构和生活方式也发生了很大改变,高血压等慢性病的患病率逐年上升,而且呈低龄化趋势,这些慢性病及其并发症造成严重的社会负担和经济负担。尽早识别个体患高血压的风险并对其进行针对性的疾病预防,在促进个体健康和减少医疗护理负担方面格外重要。

[0003] 然而,现有基于眼底彩照进行高血压风险预测的方法,单纯的以动静脉血管进行分割与识别,研究因素单一,缺乏对眼底出血、硬性渗出、视乳头水肿等其他眼底病变进行分析,也没有对个体的体格与生活方式进行分析,从而导致形成的高血压风险预测准确率不高。

发明内容

[0004] 本发明提供一种高血压风险预测方法、装置、设备及介质,其主要目的在于通过深度学习分别构建动静脉分割模型与视盘分割模型,从而在预设的视频范围内分割并识别动静脉血管,量化处理动静脉血管得到动脉、静脉的血管直径量化值;基于深度学习方法构建病变识别模型识别眼底病变特征,结合血管直径量化值、病变特征、其他风险因子等多个变量,基于机器学习回归方法建立预测待测目标高血压的风险预测模型,通过风险预测模型提高了高血压风险预测的准确性。

[0005] 为实现上述目的,本发明提供一种高血压风险预测方法,该方法包括:

[0006] 获取待测目标的个人信息与眼底图像,所述个人信息至少包含待测目标的年龄、身体质量指数、吸烟史、饮酒史、收缩压或舒张压中的一种;

[0007] 将所述眼底图像分别输入预设的动静脉分割模型、视盘分割模型与病变识别模型,分别提取所述眼底图像中的动静脉血管、视盘区域以及眼底病变特征;

[0008] 对所述视盘区域内的动静脉血管的管径进行量化处理,得到所述动静脉血管各自对应的血管直径量化值;

[0009] 利用机器学习回归方法基于所述血管直径量化值、眼底病变特征以及个人信息构建用于预测待测目标高血压的风险预测模型;

[0010] 将待测目标所对应的眼底图像输入所述风险预测模型进行检测,得到待测目标的高血压预测结果。

[0011] 可选的,所述将所述眼底图像分别输入预设的动静脉分割模型、视盘分割模型与病变识别模型之前,还包括:

[0012] 采用UNET网络基于训练集分别构建动静脉分割模型和视盘分割模型,所述病变识别模型根据不同的病变症状采用不同的神经网络模型进行训练,其中,所述病变症状至少包括动静脉压迫征、视乳头水肿、出血和渗出。

[0013] 可选的,所述采用UNET网络基于训练集分别构建动静脉分割模型和视盘分割模型,包括:

[0014] 基于训练集采用两个UNET网络训练得到动静脉分割模型和视盘分割模型;所述动静脉分割模型中一个UNET网络用于提取所述眼底图像内的血管组织特征,另一个UNET网络用于根据所述血管组织特征输出概率确定所述血管组织为动脉或静脉;所述视盘分割模型中一个UNET网络用于提取所述眼底图像内的视盘组织特征,另一个UNET网络用于根据所述视盘组织特征输出概率确定所述视盘组织是否为视盘;;

[0015] 其中,所述UNET网络包括编码层与解码层,所述编码层由卷积层与下采样层构成,用于提取特征;所述解码层由卷积层与下采样层构成,用于识别提取的特征;用于提取特征的UNET网络的损失函数为Dice Loss;用于识别特征的UNET网络的损失函数为多类交叉熵损失。

[0016] 可选的,所述病变识别模型根据不同的病变症状采用不同的神经网络模型进行训练,还包括:

[0017] 将所述待测目标的眼底图像按照病变症状不同分别进行标注,分别构成训练不同神经网络模型的多个数据集,多个所述数据集至少包括第一数据集、第二数据集、第三数据集与第四数据集;

[0018] 基于所述第一数据集内标注的动静脉压迫征病变区域和非动静脉压迫征病变区域,采用所述第一数据集训练卷积神经网络得到动静脉压迫征识别模型;

[0019] 基于所述第二数据集内标注的动静脉压迫征病变区域和非动静脉压迫征病变区域,采用所述第二数据集训练卷积神经网络得到视乳头水肿识别模型;

[0020] 基于所述第三数据集内标注的出血病变区域和非出血病变区域第三训练数据,采用所述第三数据集训练全卷积深度学习网络得到出血病变识别模型;

[0021] 基于所述第四数据集内标注的渗出病变区域和非渗出病变区域,采用所述第四数据集训练全卷积深度学习网络得到动静脉压迫征识别模型。

[0022] 可选的,所述将所述眼底图像分别输入预设的动静脉分割模型、视盘分割模型与病变识别模型之前,包括:

[0023] 将眼底图像转换成预设规格大小的图像;

[0024] 对预设规格的所述眼底图像进行归一化处理,得到像素色彩均匀的眼底图像;

[0025] 对所述像素色彩均匀的眼底图像分别进行亮度、对比度、色彩的增强处理,得到增强后的眼底图像。

[0026] 可选的,所述对所述视盘区域内的动静脉血管的管径进行量化处理,得到所述动静脉血管各自对应的血管直径量化值,还包括:

[0027] 提取预设视盘直径范围内的动脉血管与静脉血管,采用修正的Parr-Hubbard公式分别计算视网膜中央动脉直径等效值和视网膜中央静脉直径等效值,得到动脉血管与静脉血管各自对应的血管直径量化值;

$$[0028] \quad CRAE = \sqrt{0.88 \times (A_i^2 + A_j^2)}$$

$$[0029] \quad CRVE = \sqrt{0.95 \times (V_i^2 + V_j^2)}$$

[0030] 其中, A_i 和 A_j 分别为迭代获取到的最大动脉血管直径和最小动脉血管直径; V_i 和 V_j 分别为迭代获取到的最大静脉血管直径和最小静脉血管直径, CRAE 和 CRVE 分别为视网膜中央动脉直径等效值和视网膜中央静脉直径等效值。

[0031] 可选的, 所述利用机器学习回归方法基于所述血管直径量化值、眼底病变特征以及个人信息构建用于预测待测目标高血压的风险预测模型, 还包括:

[0032] 利用回归分析逐一分析所述血管直径量化值、眼底病变特征以及个人信息, 得到相关性满足预设值的变量因子;

[0033] 根据各个变量因子与高血压的相关性, 按照主成分分析法确定各个变量因子的权重系数;

[0034] 将所述变量因子以及对应的权重系数输入基于机器学习的多因素回归模型进行训练, 得到预测待测目标高血压的风险预测模型。

[0035] 此外, 为实现上述目的, 本发明还提供一种高血压风险预测装置, 所述装置包括:

[0036] 获取模块, 用于获取待测目标的个人信息与眼底图像, 所述个人信息至少包含待测目标的年龄、身体质量指数、吸烟史、饮酒史、收缩压或舒张压中的一种;

[0037] 特征提取模块, 用于将所述眼底图像分别输入预设的动静脉分割模型、视盘分割模型与病变识别模型, 分别提取所述眼底图像中的动静脉血管、视盘区域以及眼底病变特征;

[0038] 血管直径量化模块, 用于对所述视盘区域内的动静脉血管的管径进行量化处理, 得到所述动静脉血管各自对应的血管直径量化值;

[0039] 风险预测模型构建模块, 利用机器学习回归方法基于所述血管直径量化值、眼底病变特征以及个人信息构建用于预测待测目标高血压的风险预测模型;

[0040] 风险预测模块, 用于将待测目标所对应的眼底图像输入所述风险预测模型进行检测, 得到待测目标的高血压预测结果。

[0041] 此外, 为实现上述目的, 本发明还提供一种计算机设备, 包括存储器、处理器及存储在存储器上并可在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时实现如上任一项实施例所述方法的步骤。

[0042] 此外, 为实现上述目的, 本发明还提供一种计算机可读存储介质, 其上存储有计算机程序, 所述计算机程序被处理器执行时实现如上任一项实施例所述方法的步骤。

[0043] 本发明提出的高血压风险预测方法、装置、设备及介质, 该方法通过深度学习分别构建动静脉分割模型与视盘分割模型, 从而在预设的视频范围内分割并识别动静脉血管, 量化处理动静脉血管得到动脉、静脉的血管直径量化值; 基于深度学习方法构建病变识别模型识别眼底病变特征, 结合血管直径量化值、病变特征、以及个人信息生活方式相关的其他风险因子, 基于机器学习回归方法建立预测待测目标高血压的风险预测模型, 由于风险预测模型采用多个变量因子进行训练, 相对于因素单一模型构建方式, 大大提高了高血压风险预测的准确性。

附图说明

- [0044] 图1为本发明一个实施例中提供的高血压风险预测方法的一种流程示意图；
- [0045] 图2为本发明一个实施例中提供的高血压风险预测方法中眼底图像预处理流程示意图；
- [0046] 图3为本发明一个实施例中提供的高血压风险预测方法中风险预测模型构建流程示意图；
- [0047] 图4为本发明一个实施例中提供的高血压风险预测装置的一种结构示意图；
- [0048] 图5为本发明一个实施例中提供的高血压风险预测装置的一种完整结构示意图；
- [0049] 图6为发明一个实施例中提供的计算机设备的一种结构示意图。
- [0050] 本发明目的的实现、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

- [0051] 应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。
- [0052] 为了便于理解本申请,首先对本申请涉及的概念进行解释。
- [0053] 高血压是常见的慢性疾病之一,也是心脑血管病最主要的危险因素。截至2015年,我国成年人人群中高血压的患病率为27.9%且呈上升趋势。高血压会增加心、脑、肾及其他疾病的风险,尽早识别个体患高血压的风险并对其进行针对性的疾病预防,在促进个体健康和减少医疗护理负担方面格外重要。
- [0054] 眼底彩照是一种无创的非介入的成像方式,可以直接观察到眼底微血管系统的形态变化,同时,血压升高也会造成血管壁压力增加,可能导致视网膜动脉功能性狭窄、动脉硬化、动静脉交叉压迹等视网膜血管异常。通过前瞻性研究表明,视网膜病变、视网膜管径变化,尤其是动静脉管径比值变小与高血压的发生显著相关。
- [0055] 在一个实施例中,提供一种高血压风险预测方法,参照图1所示,该方法包括以下步骤:
- [0056] S101,获取待测目标的个人信息与眼底图像,所述个人信息至少包含待测目标的年龄、身体质量指数、吸烟史、饮酒史、收缩压或舒张压中的一种。
- [0057] 针对步骤S101,主要体现的是待测目标提供相应的个人信息与眼底图像,其中,个人信息报包括但不限于年龄、BMI(体质指数)、吸烟史、饮酒史、收缩压和舒张压等信息,形成影响高血压预测分析的变量因子。
- [0058] 其中,眼底图像是指眼球内后部的组织,可以包括眼球的内膜、视网膜、黄斑和血管(视网膜动静脉)。在一些示例中,可以通过监测人眼的眼底的血管变化对病变特征进行识别,其中,待测目标的个人信息与眼底图像是相互关联的,即每个待测目标的个人信息对应绑定相应的眼底图像。
- [0059] 例如,选取某一医院的眼底图像与个人信息形成数据集,按照个人信息对应的眼底彩照数据。其中各病种数据分别为:正常、出血灶、硬性渗出、动静脉压迫征、视乳头水肿,例如,训练病变识别模型采用出血灶、硬性渗出、动静脉压迫征、视乳头水肿的相应标记数据;训练动静脉分割模型采用动脉与静脉各自的分割图像,训练视盘分割模型采用眼底中的视盘区域分割图像。
- [0060] 其中,视盘区域在眼底图像上以方框、圆、椭圆、不规则多边形等方式勾勒出需要

处理的区域。

[0061] 在一些实施例中,为了得到一个较好的训练集与测试集,需要对待测目标的眼底图像进行预处理,请参阅图2,该方法还包括:

[0062] 步骤S201,将眼底图像转换成预设规格大小的图像;

[0063] 步骤S202,对预设规格的所述眼底图像进行归一化处理,得到像素色彩均匀的眼底图像;

[0064] 步骤S203,对所述像素色彩均匀的眼底图像进行数据分别进行亮度、对比度、色彩的增强处理,得到增强后的眼底图像。

[0065] 其中,将图像统一缩放到 512×512 的大小;进行归一化的操作,将图像高斯模糊处理后与原始图像反向叠加,并将像素色彩均值移动到128;对图像进行数据增强操作,对每一张训练样本进行以下操作:以百分之五十的概率进行水平翻转;以百分之五十的概率进行转置;随机gamma变换,gamma值限制在(50,150)之间;以百分之五十的概率随机改变图片的HUE、饱和度和值;以百分之五十的概率进行平移旋转操作;在以下组合中任选一项进行操作:限制对比度直方图均衡(CLAHE)算法;图像锐化后与自身重叠(IAASharpen)的算法;对图像执行某一程度浮雕操作,通过某一通道将结果与图像融合(IAAEmboss)的算法;随机亮度与对比度调节;随机模糊与增加高斯噪声处理。

[0066] 通过上述方式,一方面,通过预处理方式,能够将格式不同、大小不同的眼底图像转换为预设格式,同时,对预设规格的眼底图像进行归一化处理,得到像素色彩均匀的眼底图像;另一方面,对像素色彩均匀的眼底图像进行增强处理,能够得到亮度、清晰度、饱和度值达标的眼底图像。

[0067] 步骤S102:将所述眼底图像分别输入预设的动静脉分割模型、视盘分割模型与病变识别模型,分别提取所述眼底图像中的动静脉血管、视盘区域以及眼底病变特征。

[0068] 针对步骤S102,主要体现的是分别训练动静脉分割模型、视盘分割模型以及病变识别模型,依次提取眼底图像中的动静脉血管、视盘区域以及眼底病变特征。

[0069] 其中,采用UNET网络基于训练集分别构建动静脉分割模型和视盘分割模型,所述病变识别模型根据不同的病变症状采用不同的神经网络模型进行训练,其中,所述病变症状至少包括动静脉压迫征、视乳头水肿、出血和渗出。

[0070] 可选的,所述采用UNET网络基于训练集分别构建动静脉分割模型和视盘分割模型的步骤,包括:

[0071] 步骤S301,基于训练集采用两个UNET网络通过训练得到动静脉分割模型和视盘分割模型;

[0072] 步骤S302,所述动静脉分割模型中一个UNET网络用于提取所述眼底图像内的血管组织特征,另一个UNET网络用于根据所述血管组织特征输出概率确定所述血管组织为动脉或静脉;所述视盘分割模型中一个UNET网络用于提取所述眼底图像内的视盘组织特征,另一个UNET网络用于根据所述视盘组织特征输出概率确定所述视盘组织是否为视盘;;

[0073] 其中,所述UNET网络包括编码层与解码层,所述编码层由卷积层与下采样层构成,用于提取特征;所述解码层由卷积层与下采样层构成,用于识别提取的特征;用于提取特征的UNET网络的损失函数为Dice Loss;用于识别特征的UNET网络的损失函数为多类交叉熵损失。

[0074] 在一些实施例中,利用所述动静脉分割模型提取并识别动静脉的方法,包括:

[0075] 将眼底图像输入所述动静脉分割模型中的第一UNET网络,得到分割的血管特征;

[0076] 将所述血管特征输入所述动静脉分割模型中的第二UNET网络,根据输出概率确定动静脉分类结果。

[0077] 其中,所述UNET网络包括编码层与解码层,所述编码层由卷积层与下采样层构成,用于提取特征;所述解码层由卷积层与下采样层构成,用于识别提取的特征;用于提取特征的UNET网络的损失函数为Dice Loss;用于识别特征的UNET网络的损失函数为多类交叉熵损失。

[0078] 在本实施例中,动静脉分割方法的深度学习网络采用了血管提取深度学习网络与动静脉识别深度学习网络的两阶段,即,先用血管提取深度学习网络进行血管提取,得到血管分割图,再将血管分割图输入动静脉识别深度学习网络进行动静脉识别。其中,采用了多阶段的训练策略,能够很好的自动区分动脉、静脉,提高了动静脉分割识别的准确率,具有极高的鲁棒性。

[0079] 例如,将眼底图像输入第一UNET网络,得到关于血管组织的二值化的图像;利用第二UNET网络将该图像与眼底图像进行卷积,得到动静脉的分类结果。

[0080] 可选的,UNET网络具有下采样过程与上采样过程,且在相同层之间具有一个连接模块,卷积层之间的连接用残差模块制造恒等映射,使网络结构朝着恒等映射的方向去收敛;采用多层次输入和多尺度特征结合的多层级特征结合,且引入空洞卷积层;在下采样过程中,采用跨步为2的卷积操作;在上采样过程中使用反卷积操作;经过一个自注意力模型的掩膜层,掩膜的输入为 $h \times w \times c$ 掩膜层分别在通道方向 c 和其垂直的平面 $h \times w$ 进行卷积操作,都使用sigmoid激活函数并两者相乘,得到数值范围在 $0 \sim 1$,尺寸为 $h \times w \times c$ 的概率输出,将输入和掩膜输出相乘得到输出;掩膜输入与输出相乘分别经过卷积网络得到特征分割图,如,动静脉分割图。

[0081] 可选的,第一Unet网络的损失函数为Dice Loss具体为:

$$[0082] \quad L_s = 1 - \frac{2|x \cap y| + 1}{|x| + |y| + 1} \quad \text{公式 (1)}$$

[0083] 其中,式(1)中 $|x \cap y|$ 表示眼底图像(即,样本) x 和 y 之间的交集, $|x|$ 和 $|y|$ 分别表示 X 和 Y 的元素的个数,分子的系数为2,是因为分母存在重复计算 x 和 y 之间的共同元素的原因;另外,在分子与分母都加1,一方面,避免当 $|x|$ 和 $|y|$ 都为0时,分子被0除的问题;另一方面,减少过度拟合,通过上述方式,能够准确提取到眼底图像中的血管组织。

[0084] 可选的,第二Unet网络的损失函数为多类交叉熵损失具体为:

$$[0085] \quad L(y, y', y_1, y_1') = -(y \ln y' + (1-y) \ln(1-y')) - (y_1 \ln y_1' + (1-y_1) \ln(1-y_1')) \quad \text{公式 (2)}$$

[0086] 其中,式(2)中 y 表示眼底图像 x 真实动静脉标注概率图的概率, y_1 表示眼底图像 x 动静脉边界的真实标注概率图的概率, y' 表示眼底图像 x 真实动静脉标注概率图经 $\text{Loc}(\cdot)$ 的输出, y_1' 表示眼底图像 x 动静脉边界的真实标注概率图经 $\text{Loc}(\cdot)$ 的输出。

[0087] 通过上述多类交叉熵损失方式,一方面,多分类交叉熵损失求导更简单,损失仅与正确类别的概率有关;另一方面,可以确保提取的动静脉血管的准确性。

[0088] 在一些实施例中,利用所述视盘分割模型提取并识别视盘的方法,包括:

[0089] 将眼底图像输入所述视盘分割模型中的第一Unet网络,得到视盘特征;

[0090] 将所述视盘特征输入所述视盘分割模型中的第二Unet网络,根据输出概率确定视盘分类结果。

[0091] 需要说明的是,由于视盘分割模型的提取并识别方式与动静脉分割模型的提取并识别方式相同,在此不再赘述。

[0092] 可选的,所述病变识别模型根据不同的病变症状采用不同的神经网络模型进行训练的步骤,还包括:

[0093] 将所述待测目标的眼底图像按照病变症状不同分别进行标注,分别构成训练不同神经网络模型的多个数据集,多个所述数据集至少包括第一数据集、第二数据集、第三数据集与第四数据集;

[0094] 基于所述第一数据集内标注的动静脉压迫征病变区域和非动静脉压迫征病变区域,采用所述第一数据集训练卷积神经网络得到动静脉压迫征识别模型;

[0095] 基于所述第二数据集内标注的动静脉压迫征病变区域和非动静脉压迫征病变区域,采用所述第二数据集训练卷积神经网络得到视乳头水肿识别模型;

[0096] 基于所述第三数据集内标注的出血病变区域和非出血病变区域第三训练数据,采用所述第三数据集训练全卷积深度学习网络得到出血病变识别模型;

[0097] 基于所述第四数据集内标注的渗出病变区域和非渗出病变区域,采用所述第四数据集训练全卷积深度学习网络得到动静脉压迫征识别模型。

[0098] 在一些实施例中,利用病变识别模型识别眼底病变特征的步骤,具体包括:

[0099] 将同一待识的别眼底图像分别导入动静脉压迫征识别模型、视乳头水肿识别模型、出血病变识别模型和渗出病变识别模型中进行识别;

[0100] 根据识别结果提取病变特征信息,再采用经训练的支持向量机分类器对提取的病变特征信息分类,获得眼底图像对应的病变类型分类结果;

[0101] 所述动静脉压迫征识别模型通过提取眼底图像中动静脉病变候选区域,进行动静脉压迫征病变区域和非动静脉压迫征病变区域标注,再输入CNN模型进行训练得到;所述视乳头水肿识别模型通过提取眼底图像中视乳头水肿候选区域,进行视乳头水肿区域和视乳头非水肿区域标注,再输入CNN模型进行训练得到;所述出血病变识别模型通过对眼底图像中出血病变区域和非出血病变区域进行标注,再输入FCN模型进行训练得到;所述渗出病变识别模型通过对眼底图像中渗出病变区域和非渗出病变区域进行标注,再输入FCN模型进行训练得到。

[0102] 具体地,病变特征信息包括但不限于病变区域的个数、面积、形状、灰度、圆度和纵横比。

[0103] 在本实施例中,通过对不同的病变症状采取了不同的神经网络模型,因此,保存的模型针对特定的病变识别具有更高的精度和更强的适用性,综合动静脉压迫征、视乳头水肿、出血、渗出四种病变的多个特征进行分类,具有更高的分类准确率,可以更加有效地辅助医生进行临床筛查工作。

[0104] 在另一些实施例中,所述出血病变识别模型和渗出病变识别模型均基于FCN模型训练得到,包括以下步骤:

[0105] 通过图像处理对眼底图像进行病变区域和非病变区域标注,生成训练集;使用UNET网络结构构建FCN模型,每次随机取训练集中部分已标注数据进行训练,得到训练后的病变识别模型;训练采用的损失函数是DICE。

[0106] 其中,FCN模型为全卷积神经网络,包括卷极化、上采样以及跳跃结构,其中,卷极化,用于适应任意尺寸输入,输出低分辨率的分割图像;上采样(反卷积),用于在低分辨率的图像进行上采样,输出同分辨率的分割图像;跳跃结构,用于结合上采样和上层卷积池化后的数据,修复还原图像,在实施例中涉及的图像指眼底图像。

[0107] 具体地,对眼底图像的标注还生成测试集,对已训练模型采用测试集测试,评估已训练模型的识别能力;对眼底图像的标注还生成验证集,在模型训练中采用验证集进行修正,防止网络过拟合。

[0108] 步骤S103:对所述视盘区域内的动静脉血管的管径进行量化处理,得到所述动静脉血管各自对应的血管直径量化值;

[0109] 针对步骤S103,主要体现的是如何将步骤S102中分类的动脉、静脉转化为与高血压变量因子相关的系数,经医学证明,患有高血压用户与未患有高血压的用户在眼底图像中微血管明显存在差异,如:在反复的高压刺激下,高血压中视网膜小动脉早期会稍有狭窄和轻度硬化;如果血压长期增高,发展到一定阶段,视网膜会进一步发生改变,动脉持续狭窄,视网膜动脉硬化明显,动脉会出现“银线反应”,动脉管径狭窄不均,并有动静脉交叉压迫现象;从上可知,高血压能够影响或降低动静脉管的直径。

[0110] 具体地,定位眼底图像对应的预设视盘直径范围内动脉管与静脉管,采用Parr-Hubbard-knudtson公式分别计算CRAE(centralretinalarteryequivalent,视网膜中央动脉直径等效值)和CRVE(centralretinalveinequivalent,视网膜中央静脉直径等效值),Parr-Hubbard-knudtson公式为:

$$[0111] \quad CRAE = \sqrt{0.88 \times (A_i^2 + A_j^2)} \quad (\text{公式 } 3)$$

$$[0112] \quad CRVE = \sqrt{0.95 \times (V_i^2 + V_j^2)} \quad (\text{公式 } 4)$$

[0113] 其中,式(3)与式(4)中 A_i 和 A_j 分别为迭代获取到的最大动脉血管直径和最小动脉血管直径; V_i 和 V_j 分别为迭代获取到的最大静脉血管直径和最小静脉血管直径,,最后,通过公式 $AVR = CRAE/CRVE$ 获得最终的AVR定量计算结果。

[0114] 通过上述方式,能够将视网膜的动脉血管与静脉血管分别进行量化,得到视网膜中央动脉直径等效值与视网膜中央静脉直径等效值,进而得到动脉血管与静脉血管各自对应的血管直径量化值。

[0115] 步骤S104:利用机器学习回归方法基于所述血管直径量化值、眼底病变特征以及个人信息构建用于预测待测目标高血压的风险预测模型。

[0116] 针对步骤S104,主要体现的是为了构建一个更加全面的高血压预测模型,包括但不限于待测目标的生活习惯(饮酒、吸烟等)、待测目标的BMI、年龄等个人信息;还涉及血管量化值的血管管径、以及视网膜病变等变量因子,从全方位、多角度考虑各种变量因子,能够更细微的反应高血压的预测等级,也能提高-高血压的预测精准度。

[0117] 在一些实施例中,基于所述血管量化值、眼底病变特征以及个人信息构建高血压预测模型的步骤包括:

[0118] 步骤S401,利用回归分析逐一分析所述血管直径量化值、眼底病变特征以及个人信息,得到相关性满足预设值的变量因子;

[0119] 步骤S402,根据各个变量因子与高血压的相关性,按照主成分分析法确定各个变量因子的权重系数;

[0120] 步骤S403,将所述变量因子以及对应的权重系数输入基于机器学习的多因素回归模型进行训练,得到预测待测目标高血压的风险预测模型。

[0121] 具体地,血管直径量化值(CRAE、CRVE、AVR)、病变征象有无(出血灶、硬性渗出、动静脉压迫征、视乳头水肿)、其他风险因子(年龄、BMI、吸烟史、饮酒史)与高血压的相关性,首先,对每个单因素进行回归分析,选择显著相关(P值小于0.05)的变量纳入到多因素回归模型中,将上一步中筛选出来的相关变量输入多因素回归模型进行高血压风险预测建模,从而实现高血压风险等级预测的功能,建模公式如下:

$$[0122] \quad \text{Risk}_{\text{高血压}} = \frac{1}{1 + e^{-(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}} \quad (\text{公式 } 5)$$

[0123] 其中,公式(5)中 x_1 、 \dots 、 x_n 为上述变量因子, β_1 、 \dots 、 β_n 为各变量因子的权重,该值的大小表明各变量因子每增加1个单位,其患病风险增至 e^β 倍。

[0124] 通过上述方式,筛选多个高血压呈正相关的风险因子,构建了一个高精度的高血压预测模型,大大提升了待测目标的高血压的预测精准度。

[0125] 在另一些实施例中,还包括:基于统计学方法筛选确定高血压预测模型的变量因子,确定影响高血压患病的变量,根据二元Logistic回归方法进行相关性分析,得到各个变量的隐状态P值,选择P值小于设定阈值的变量作为变量因子。

[0126] 具体实施中,可以首先依据《中国高血压健康管理规范(2019)》及近年来的高血压人群统计表,首先确定了一些可能影响高血压患病的因素,即变量。包括15个:年龄、性别、吸烟、运动、高血压家族史、肥胖、糖尿病、长期精神紧张、吸烟、高血脂、高盐摄入、收缩压、舒张压、过量饮酒和空气污染筛选目的是在15个变量中,排除掉预测效能差的变量,筛选出强相关变量,作为后续预测模型建立的基础。

[0127] 依据数据库提供的健康信息数据,取出高血压患病因素数据,采用统计学方法对变量进行筛选,依据二元Logistic回归方法进行相关性分析,得到单个变量的P值;依据P值小于0.05具有统计学意义,故而排除了对高血压影响较小的因素。最终筛选的变量因子包括:年龄、性别、吸烟、运动、高血压家族史、肥胖(BMI)、糖尿病、收缩压和舒张压。选择的9个变量因子在单因素分析中全部达到筛选条件($P < 0.05$)。其中肥胖用BMI来表示, $BMI = \text{体重(千克)} / \text{身高(平方米)}$ 。

[0128] 在上述变量因子的基础上,结合血管直径量化值、病变特征(出血灶、硬性渗出、动静脉压迫征、视乳头水肿)作为强相关的变量因子,确定各个变量因子的回归系数及对应分值。

[0129] 根据各个变量因子的回归系数及对应分值构建高血压预测的多因素Logistic回归模型,多因素Logistic回归模型为各个变量因子对应的高血压风险预测概率值函数。

[0130] 利用公式(5),通过构建多因素Logistic回归模型,将主要考虑的危险因素纳入到

该多因素Logistic回归模型中,从而估计各个因素的回归系数 β 、OR (Odds ratio, 胜率) 值及其95%CI (Confidence interval, 置信区间)。多因素Logistic回归模型中,OR值为1,表示该因素对疾病的发生不起作用;OR值大于1,表示该因素为一个危险因素;OR值小于1,表示该因素是一个保护因素。如果回归系数 β 为正,表明随着自变量的增加,因变量的对数即 $\ln(p/1-p)$ 也在变大,必然因变量取值的概率 p 也在增加,但这时增加的是因变量取低值的概率,那么该自变量应该和因变量较小的取值相联系;反之,回归系数 β 为负,说明该自变量对因变量的影响是负的,也就是负相关。

[0131] 在一些实施例中,确定各个变量因子对应参考值的方法包括:对各个变量因子的值进行分组。

[0132] 变量因子为数值变量时,根据变量因子的数值范围设置各个分段范围的分组,每个分组中选择中间值为参考值 M_{ref} 。

[0133] 变量因子为分类变量时,根据变量因子的类型设置各个为两个分组,两个分组的参考值 M_{ref} 为0或1。

[0134] 变量因子为数值变量时,按照临床意义或使用习惯将各个危险因素进行分组,并在每个分组中选择合适的数值作为参考值 M_{ref} ,通常选择组内的中间值作为参考值。

[0135] 例如,在本实施例中,研究人群的年龄范围为45-94岁,通常我们按照10岁一个年龄段来进行划分,将其分为5组,每组选择中间值为参考值 M_{ref} ,例如45-54岁这一组的参考值 M_{ref} 为 $(45+54)/2=49.5$ 。

[0136] 收缩压的范围为70-139mmHg, <110mmHg为一组,110mmHg以上的我们按照每5mmHg为一组,共分为7组,每组选择中间值为参考值 M_{ref} 。例如120-124mmHg这一组的参考值 M_{ref} 为 $(120+124)/2=122$ 。

[0137] 舒张压的范围为50-89mmHg, <70mmHg为一组,70mmHg以上的我们按照每10mmHg为一组,共分为3组,每组选择中间值为参考值 M_{ref} 。例如70-80mmHg这一组的参考值 M_{ref} 为 $(70+80)/2=75$ 。

[0138] BMI的范围为15-50, <25为一组,25-29为一组,30-39为一组, ≥ 40 的为一组,每组选择中间值为参考值 M_{ref} 。例如25-29这一组的参考值 M_{ref} 为 $(25+29)/2=27$ 。

[0139] 变量因子为分类变量时,如性别,此时可以设置男性为参照,即参考值 M_{ref} 为0,那么女性就自然赋值为1,同理,不吸烟设置为0,吸烟为1;运动设置为0,不运动为1;无高血压家族史设置为0,高血压家族史设置为1;无糖尿病设置为0,糖尿病设置为1。

[0140] 血管直径量化值(CRAE、CRVE、AVR),例如:随着动脉血压的升高视网膜血管的直径减少,CRAE、CRVE、AVR均与血压相关MABP每升高10mmHg,CRAE、CRVE、AVR分别减少3.5微米、0.96微米和0.012,由于血管直径量化按照1~6进行计算,每个量化指数对应2,其对应的参考值 M_{ref} 分值为3。

[0141] 病变(出血灶、硬性渗出、动静脉压迫征、视乳头水肿)症状,动静脉交叉压迫征、出血灶、硬性渗出、视乳头水肿OR值为95%,CI分别为1.25(1.16~1.34)、1.25(1.15~1.37)、1.25(1.16~1.36)和1.25(1.13~1.38),由于病变症状每出现一项,分数递增2分。

[0142] 第 i 个变量因子的对应分值Point的确定方法包括:

[0143] 选择作为基础风险参考值 M_{ref} 的变量因子的分组。

[0144] 对于每一个危险因素,需要选择一个合适的分组作为风险参考值 M_{ref} ,在构建多因

素Logistic回归模型时,该组分值将记为0分,危险因素的值高于 R_{ref} 时记正分,得分越高则风险越高,相反低于 R_{ref} 时记负分。

[0145] 例如,可以选择年龄45-54岁、男性、不吸烟、运动、 $BMI < 25$ 、无高血压家族史、无糖尿病、收缩压 $< 110\text{mmHg}$ 、舒张压 $< 70\text{mmHg}$ 对应的参考值 M_{ref} ,作为每个危险因素的基础风险参考值 R_{ref} 。

[0146] 结合回归系数 β_i ,计算各个变量的分组与基础风险参考值 R_{ref} 之间的距离 $D = (M_{ref} - R_{ref}) * \beta_i$ 。

[0147] 例如在本发明实施例中,年龄的基础风险参考值 R_{ref} 为44.5,年龄在多因素Logistic回归模型中对应的回归系数 β_i 为0.0575,那么对于55-64岁组,其参考值 M_{ref} 为59.5,该组与基础风险参考值的距离即为 $(59.5 - 44.5) * 0.0575 = 0.8625$ 。同理,其他危险因素也按照上述公式计算每个分组到基础风险参考值的距离 D 。

[0148] 确定常数 $B = x * \beta_i$, x 表示变量因子的分组的间隔。

[0149] 常数 B 为设定评分工具中每记1分时,对应的各个危险因素变化的常数。例如本发明提供的实施例中,如果设定年龄每增加5岁时记为1分,那么此时常数 $B = 5 * \beta_i = 5 * 0.0575 = 0.2875$ 。

[0150] 计算第 i 个变量因子的对应分值 $Point = D / B = (M_{ref} - R_{ref}) * \beta_i / B$ 。

[0151] 最后可以将计算出来的数值四舍五入取整,即为该分组对应的分值。

[0152] 将每个危险因素的分值相加起来计算总分,理论上每个危险因素取最低值时,可以得到总分最低值为 $0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 = 0$,同理可以得到总分最高值为 $4 + 1 + 1 + 1 + 1 + 3 + 2 + 1 + 3 + 6 + 8 = 0$,因此总分的范围为:0-44分。

[0153] 计算总分与风险预测概率的对应表。

[0154] 优选的,还包括:对高血压发病风险按照各个分值对应的概率大小进行分层,分层包括“无”、“低”、“中”、“高”。

[0155] 例如高($> 35\%$)、中($10\% - 35\%$)、低($3\% < 10\%$)与无($< 3\%$),并针对不同的分层提供个性化、专业化的健康管理方案,包括饮食、运动、体检、日常注意、预防措施、就医指征及高血压疾病健康宣教等知识。

[0156] 通过上述方式,无需传统人工方式来判断待测目标是否患有高血压,根据待测目标的眼底图像即可预测该目标是否患有高血压,进而极大简化了工作人员的工作量,提升了高血压预测的效率,也为临床高血压的早筛提供了指导性建议。对高血压发病风险按照概率大小进行分层,如“无”、“低”、“中”、“高”,并针对不同的分层提供个性化、专业化的健康管理方案。

[0157] 步骤S105,将待测目标所对应的眼底图像输入所述风险预测模型进行检测,得到待测目标的高血压预测结果。

[0158] 针对步骤S105,主要体现的是获取待测目标的眼底图像,根据眼底图像输入提前训练好的高血压预测模型进行检测,即可得到待测目标的检测结果。

[0159] 具体地,通过高血压预测结果包括“无”、“低”、“中”、“高”四个等级,进而能够数字化高血压等级。

[0160] 在另一些实施例中,从待测目标的高血压预测结果中检测到高血压风险时,生成血压健康控制方案。

[0161] 对于任一风险因素以及干预措施,若该干预措施会减少或消除该风险因素,进而阻断、延缓或者预防疾病的发生,则认为该风险因素与该干预措施之间存在关联关系,该关联关系在图中用风险因素和干预措施之间的直线表示;若该干预措施不会减少或消除该风险因素,则认为该风险因素与该干预措施之间不存在关联关系。在对用户进行高血压预警干预时,如果用户存在该风险因素,则可以根据该关联关系推荐用户采用该干预措施。例如,减少饮酒这一干预措施会减轻长期饮酒这一风险因素,因而认为减少饮酒与长期饮酒之间存在关联关系,同理,减少抽烟与长期吸烟之间也存在关联关系;在进行健康干预时,若发现用户的健康档案中存在长期饮酒或长期吸烟这一习惯,则根据该关联关系可以推荐用户采用减少饮酒或减少吸烟这一干预措施。

[0162] 本实施例提供了一种高血压风险预测方法,该方法通过深度学习分别构建动静脉分割模型与视盘分割模型,从而在预设的视频范围内分割并识别动静脉血管,量化处理动静脉血管得到动脉、静脉的血管直径量化值;基于深度学习方法构建病变识别模型识别眼底病变特征,结合血管直径量化值、病变特征、以及个人信息生活方式相关的其他风险因子,基于机器学习回归方法建立预测待测目标高血压的风险预测模型,由于风险预测模型采用多个变量因子进行训练,相对于因素单一模型构建方式,大大提高了高血压风险预测的准确性。

[0163] 下面通过一个具体的实施例,示例性的说明上述高血压风险预测方法,该具体的高血压风险预测方法包括:

[0164] 为了进一步验证Logistic回归模型的准确性,本发明提供一个实例,来比较一下评分工具与原始Logistic回归模型预测结果之间的差距。

[0165] 假设有一位男性患者,75岁、收缩压124mmHg、舒张压87mmHg、不吸烟、运动、无高血压家族史、有糖尿病、BMI=21、血管直径量化值、病变特征(出血灶、硬性渗出、动静脉压迫征、视乳头水肿,每增加一项特征分值对应增加一项),预测他未来3年高血压的发生风险。

[0166] 首先,根据评分工具里各个危险因素的分值,分别记为0、4、6、3、0、0、0、1、0、2、4分,总分为20分,查表对应的风险概率为45%。

[0167] 然后,再根据多因素logistic回归模型进行一次计算: $y=45\%$,可以看出,评分工具与Logistic回归模型预测结果之间仅仅只相差1%,足够满足疾病风险预测评估的要求,而且应用起来也很直观和便捷。

[0168] 在一个实施例中,本发明还提供了一种高血压风险预测装置500,参见图4,该装置包括:

[0169] 获取模块501,用于获取待测目标的个人信息与眼底图像,所述个人信息至少包含待测目标的年龄、身体质量指数、吸烟史、饮酒史、收缩压或舒张压中的一种;

[0170] 特征提取模块502,用于将所述眼底图像分别输入预设的动静脉分割模型、视盘分割模型与病变识别模型,分别提取所述眼底图像中的动静脉血管、视盘区域以及眼底病变特征;

[0171] 血管直径量化模块503,用于对所述视盘区域内的动静脉血管的管径进行量化处理,得到所述动静脉血管各自对应的血管直径量化值;

[0172] 风险预测模型构建模块504,利用机器学习回归方法基于所述血管直径量化值、眼

底病变特征以及个人信息构建用于预测待测目标高血压的风险预测模型；

[0173] 风险预测模块505,用于将待测目标所对应的眼底图像输入所述风险预测模型进行检测,得到待测目标的高血压预测结果。

[0174] 本实施例中,请参阅图5,所述特征提取模块502之前,该装置还包括:

[0175] 提取模型构成模块506,采用UNET网络分别训练动静脉分割模型、视盘分割模型,所述病变识别模型根据不同的病变症状采用不同的神经网络模型进行训练,其中,所述病变症状至少包括动静脉压迫征、视乳头水肿、出血和渗出。

[0176] 本实施例中,所述提取模型构成模块进一步包括:动静脉分割模型与视盘分割模型;

[0177] 采用两个UNET网络通过训练得到动静脉分割模型、视盘分割模型;

[0178] 动静脉分割模型,一个UNET网络用于提取所述眼底图像内的血管组织特征,另一个UNET网络用于根据所述血管组织特征输出概率确定所述血管组织为动脉或静脉;和

[0179] 视盘分割模型,一个UNET网络用于提取所述眼底图像内的视盘组织特征,另一个UNET网络用于根据所述视盘组织特征输出概率确定所述视盘组织是否为视盘;

[0180] 其中,所述UNET网络包括编码层与解码层,所述编码层由卷积层与下采样层构成,用于提取特征;所述解码层由卷积层与下采样层构成,用于识别提取的特征;用于提取特征的UNET网络的损失函数为Dice Loss;用于识别特征的UNET网络的损失函数为多类交叉熵损失。

[0181] 本实施例中,所述提取模型构成模块进一步包括:病变识别模型;

[0182] 数据集构建单元,用于将所述待测目标的眼底图像按照病变症状不同分别进行标注,分别构成训练不同神经网络模型的多个数据集,多个所述数据集至少包括第一数据集、第二数据集、第三数据集与第四数据集;

[0183] 静脉压迫征识别模型,基于所述第一数据集内标注的动静脉压迫征病变区域和非动静脉压迫征病变区域,采用所述第一数据集训练卷积神经网络得到动静脉压迫征识别模型;

[0184] 视乳头水肿识别模型,基于所述第二数据集内标注的动静脉压迫征病变区域和非动静脉压迫征病变区域,采用所述第二数据集训练卷积神经网络得到视乳头水肿识别模型;

[0185] 血病变识别模型,基于所述第三数据集内标注的出血病变区域和非出血病变区域第三训练数据,采用所述第三数据集训练全卷积深度学习网络得到出血病变识别模型;

[0186] 动静脉压迫征识别模型,基于所述第四数据集内标注的渗出病变区域和非渗出病变区域,采用所述第四数据集训练全卷积深度学习网络得到动静脉压迫征识别模型。

[0187] 本实施例中,所述血管直径量化模块还包括:

[0188] 将眼底图像转换成预设规格大小的图像;

[0189] 对预设规格的所述眼底图像进行归一化处理,得到像素色彩均匀的眼底图像;

[0190] 对所述像素色彩均匀的眼底图像分别进行亮度、对比度、色彩的增强处理,得到增强后的眼底图像。

[0191] 本实施例中,所述血管直径量化模块还包括:

[0192] 提取预设视盘直径范围内的动脉血管与静脉血管,采用修正的Parr-Hubbard公式

分别计算视网膜中央动脉直径等效值和视网膜中央静脉直径等效值,得到动脉血管与静脉血管各自对应的血管直径量化值;

$$[0193] \quad CRAE = \sqrt{0.88 \times (A_i^2 + A_j^2)}$$

$$[0194] \quad CRVE = \sqrt{0.95 \times (V_i^2 + V_j^2)}$$

[0195] 其中, A_i 和 A_j 分别为迭代获取到的最大动脉血管直径和最小动脉血管直径; V_i 和 V_j 分别为迭代获取到的最大静脉血管直径和最小静脉血管直径,CRAE和CRVE分别为视网膜中央动脉直径等效值和视网膜中央静脉直径等效值。

[0196] 在本实施例中,所述风险预测模型构建模块还包括:

[0197] 利用回归分析逐一分析所述血管直径量化值、眼底病变特征以及个人信息,得到相关性满足预设值的变量因子;

[0198] 根据各个变量因子与高血压的相关性,按照主成分分析法确定各个变量因子的权重系数;

[0199] 将所述变量因子以及对应的权重系数输入基于机器学习的多因素回归模型进行训练,得到预测待测目标高血压的风险预测模型。

[0200] 本实施例提供了一种高血压风险预测装置,该装置通过深度学习分别构建动静脉分割模型与视盘分割模型,从而在预设的视频范围内分割并识别动静脉血管,量化处理动静脉血管得到动脉、静脉的血管直径量化值;基于深度学习方法构建病变识别模型识别眼底病变特征,结合血管直径量化值、病变特征、以及个人信息生活方式相关的其他风险因子,基于机器学习回归方法建立预测待测目标高血压的风险预测模型,由于风险预测模型采用多个变量因子进行训练,相对于因素单一模型构建方式,大大提高了高血压风险预测的准确性。

[0201] 应当知晓的是,上述高血压风险预测装置系统实质上是设置了多个模块用以执行上述任一实施例中的高血压风险预测方法,具体功能和技术效果参照上述实施例即可,此处不再赘述。

[0202] 在一个实施例中,参见图6,本实施例还提供了一种计算机设备600,包括存储器601、处理器602及存储在存储器上并可在处理器上运行的计算机程序,所述处理器602执行所述计算机程序时实现如上任一项实施例所述方法的步骤。

[0203] 在一个实施例中,还提供一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现如上任一项实施例所述方法的步骤。

[0204] 本申请实施例可以基于人工智能技术对相关的数据进行获取和处理。其中,人工智能(Artificial Intelligence, AI)是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。

[0205] 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、装置、物品或者方法不仅包括那些要素,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、装置、物品或者方法所固有

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、装置、物品或者方法中还存在另外的相同要素。

[0206] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。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,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明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如上所述的一个存储介质(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)执行本发明各个实施例所述的方法。

[0207] 以上仅为本发明的优选实施例,并非因此限制本发明的专利范围,凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本发明的专利保护范围内。

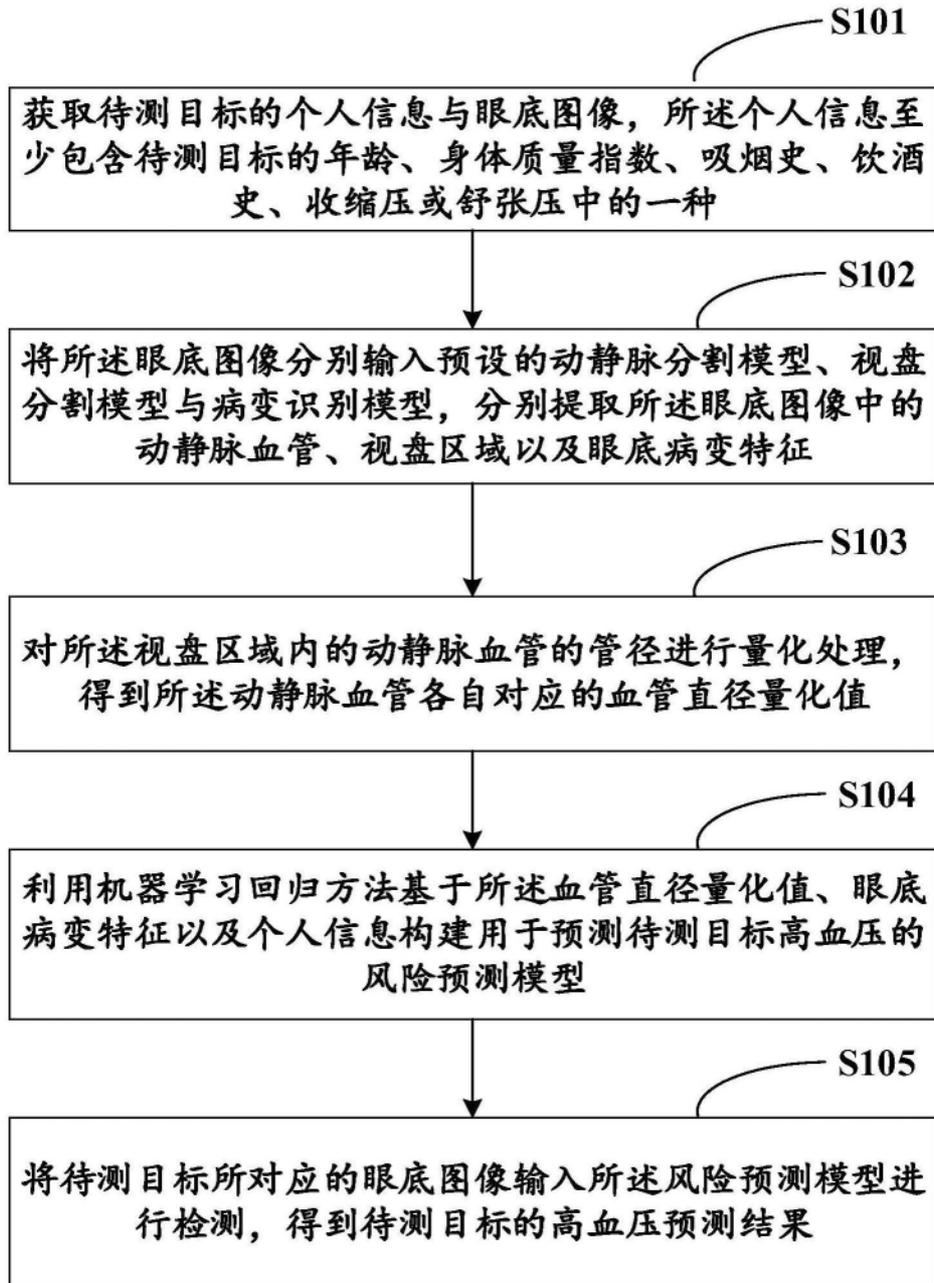


图1

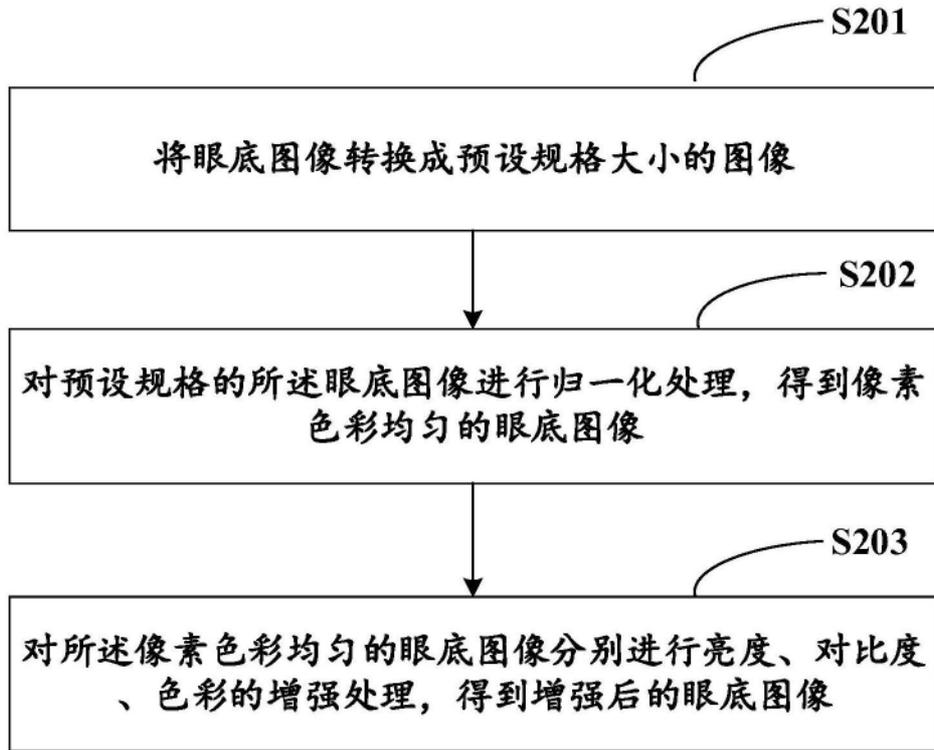


图2

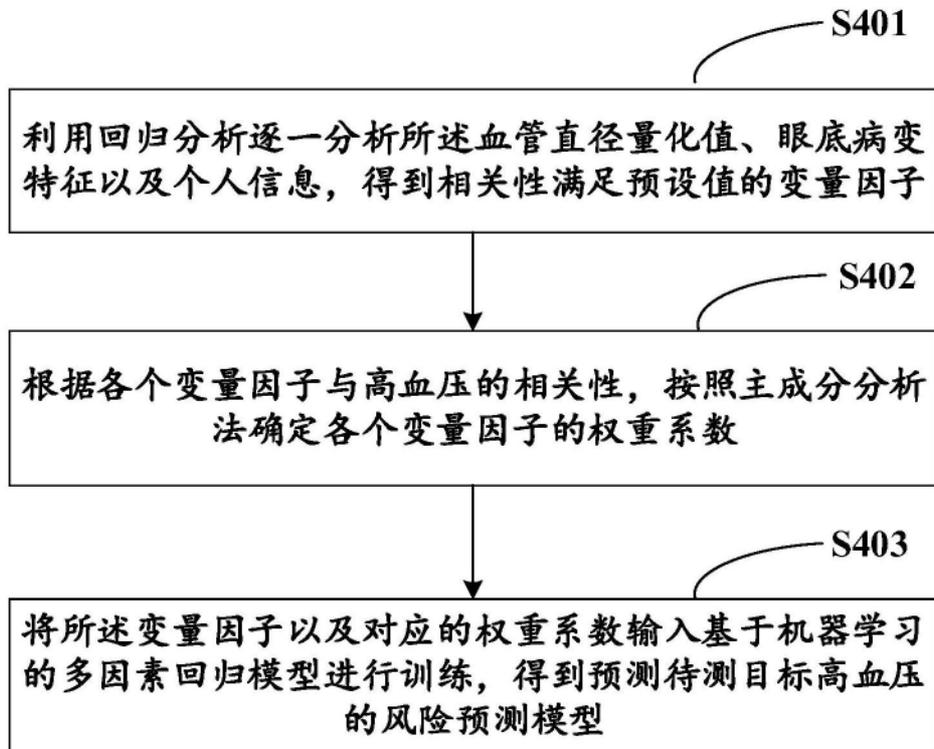


图3

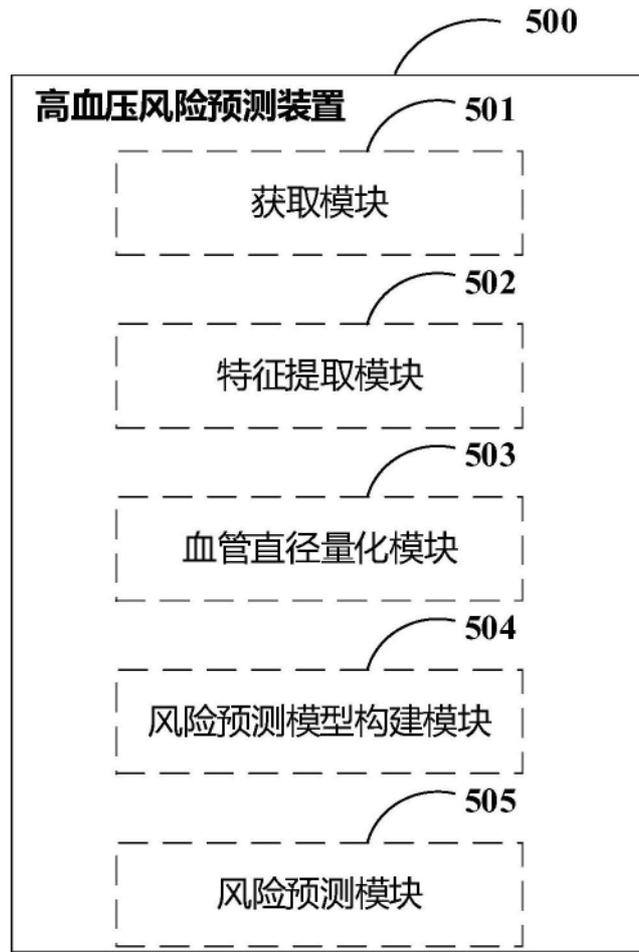


图4

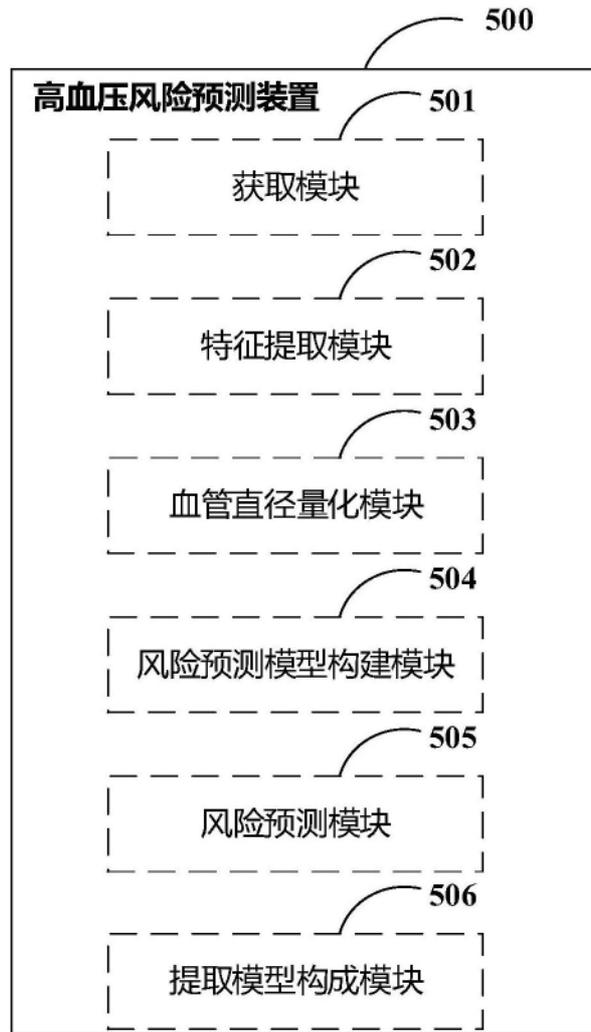


图5

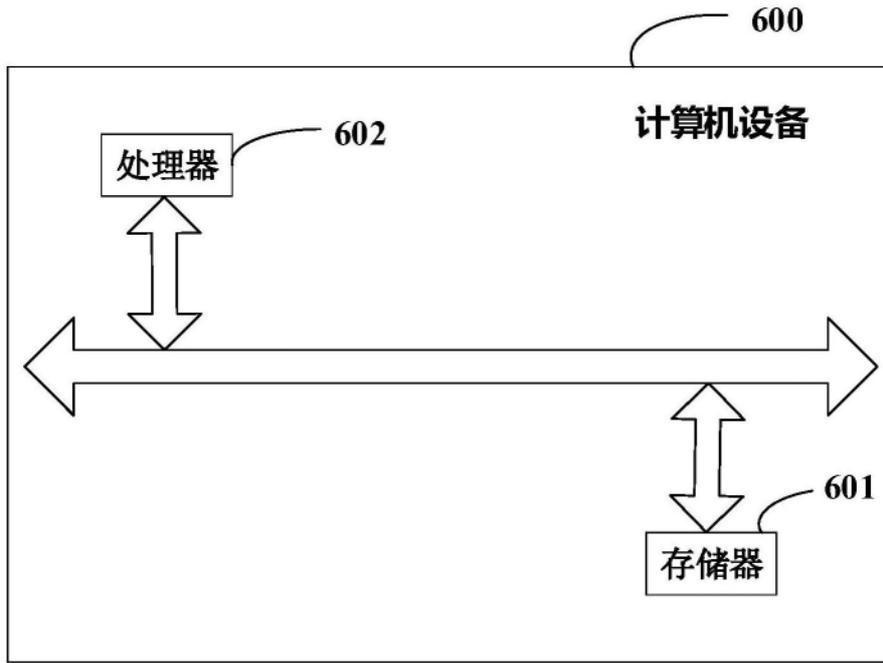


图6