

一种医学报告的生成方法及设备

本申请申明享有 2018 年 05 月 14 日递交的申请号为 201810456351.1、名称为“一种医学报告的生成方法及设备”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

5 本申请属于信息处理技术领域，尤其涉及一种医学报告的生成方法及设备。

背景技术

随着医疗影像技术的不断发展，医生可以通过医疗图像高效地确定患者的病症，诊断时长大幅减少。医生会根据医疗图像手动填写对应的医学报告，以便患者更好地获知自身的病症。但现有的医学报告的生成方法，对于患者以及见习医生而言，无法直接从医疗图像确定病症，需要依赖有经验的医生进行填写，从而增加了生成医学报告的人工成本，并且手动填写的效率也较低，无疑增加了患者的治疗时间。

技术问题

有鉴于此，本申请实施例提供了一种医学报告的生成方法及设备，以解决现有的医学报告的生成方法，生成医学报告的人工成本较高，延长患者的治疗时间的问题。

技术方案

15 本申请实施例的第一方面提供了一种医学报告的生成方法，包括：
接收待识别的医疗图像；
将所述医疗图像导入预设的视觉几何组 VGG 神经网络，得到所述医疗图像的视觉特征向量以及关键词序列；
将所述视觉特征向量以及所述关键词序列导入至预设的诊断项目识别模型，确定所述医疗图像对应的诊
20 断项目；
基于诊断项目扩展模型，分别构建用于描述各个所述诊断项目的段落；
根据所述段落、所述关键词序列以及所述诊断项目，生成所述医疗图像的医学报告。

有益效果

本申请实施例通过将医疗图像导入到预设的 VGG 神经网络，确定该医疗图像对应的视觉特征向量以及
25 关键词序列，视觉特征向量用于表征该医疗图像中包含病症的图像特征，而关键词序列则用于确定该医疗图
像中所包含的病症类型，将上述两个参数导入到诊断项目识别模型，确定该医疗图像中所包含的诊断项目，
并为每个诊断项目填充相关描述的短语以及句子，构成该诊断项目对应的段落，最后基于各个诊断项目对应
的段落得到该医疗图像的医学报告。与现有的医学报告的生成方法相比，本申请实施例无需医生手动填写，
可以自动根据医疗图像中包含的特征输出对应的医学报告，提高了医疗报告的生成效率，减少了人工成本，
30 节省了患者诊疗的时间。

附图说明

图 1a 是本申请第一实施例提供的一种医学报告的生成方法的实现流程图；

图 1b 是本申请一实施例提供的 VGG 神经网络的结构框图；

图 1c 是本申请一实施例提供的 LSTM 神经网络的结构框图；

5 图 2 是本申请第二实施例提供的一种医学报告的生成方法 S102 具体实现流程图；

图 3 是本申请第三实施例提供的一种医学报告的生成方法 S103 具体实现流程图；

图 4 是本申请第四实施例提供的一种医学报告的生成方法具体实现流程图；

图 5 是本申请第四实施例提供的一种医学报告的生成方法的具体实现流程图；

图 6 是本申请一实施例提供的一种医学报告的生成设备的结构框图；

10 图 7 是本申请另一实施例提供的一种医学报告的生成设备的示意图。

本发明的实施方式

在本申请实施例中，流程的执行主体为医学报告的生成设备。该医学报告的生成设备包括但不限于：笔记本电脑、计算机、服务器、平板电脑以及智能手机等医学报告的生成设备。图 1a 示出了本申请第一实施例提供的医学报告的生成方法的实现流程图，详述如下：

15 在 S101 中，接收待识别的医疗图像。

在本实施例中，医学报告的生成设备可以集成与医疗图像的拍摄终端内，在该情况下，拍摄终端在完成拍摄操作，生成患者的医疗图像后，则可以把该医疗图像传输给该医学报告的生成设备，对该医疗图像进行分析，确定对应的医学报告，从而无需打印医学图像给患者以及医生，从而提高了处理效率，当然医学报告的生成设备还可以只与拍摄终端的串口进行连接，通过相关的串口接口传输生成的医疗图像。

20 在本实施例中，医学报告的生成设备可以通过内设的扫描模块，对打印得到的医疗图像进行操作，从而获取计算机可读的医疗图像。当然，该生成设备还可以通过有线通信接口或无线通信接口接收用户终端发送的医疗图像，然后将分析得到的医学报告通过对应的通信信道返回给用户终端，实现远距离获取医疗报告的目的。

25 在本实施例中，医疗图像包括但不限于：各种放射光拍摄人体后的图像，如 X 光图像、B 型超声波图像等，以及病理学图像，如解剖图、基于微型导管拍摄的人体体内脏器图。

可选地，在 S101 之后，生成设备还可以通过预设的图像处理算法对医疗图像进行优化处理。上述图像处理算法包括但不限于：锐化处理、二值化处理、降噪处理、灰度处理等图像处理算法。特别地，若是通过扫描方式获取该医疗图像，则可以通过提高扫描分辨率的方式，增加获取得到的医疗图像的图像质量，并通过采集扫描时刻的环境光强，对医疗图像进行差分处理，以减少环境光对医疗图像的影响，提高后续识别的
30 准确率。

在 S102 中，将所述医疗图像导入预设的视觉几何组 VGG 神经网络，得到所述医疗图像的视觉特征向量

以及关键词序列。

在本实施例中，生成设备存储有视觉几何组（Visual Geometry Group，VGG）神经网络对医疗图像进行处理，提取该医疗图像所对应的视觉特征向量以及关键词序列。其中，视觉特征向量用于描述医疗图像中所拍摄物体的图像特征，例如轮廓特征、结构特征、各个对象之间的相对距离等；所述关键词特征用于表征该医疗图像中包含的对象以及对象的属性。例如，若医疗图像所拍摄的部位是胸部，则识别得到的关键词序列可以为：[胸、肺部、肋骨、左肺叶、右肺叶、心脏]等，当然若某个部分存在异常对象，也可以在关键词序列中体现。优选地，视觉特征向量与关键词序列各元素之间是一一对应的，即视觉特征向量中各元素是用于描述关键词序列中各关键词的图像特征。

在本实施例中，该 VGG 神经网络可以采用 VGG19 神经网络，由于 VGG19 神经网络在图像特征提取方面具有较强的运算能力，能够将包含多个图层的图像数据通过五层池化层降维运算后，提取得到视觉特征，并且在本实施例中，将全连接层调整为关键词索引表，从而能够基于关键词索引表输出关键词序列。其中，VGG19 的示意图可参见图 1b 所示。

可选地，在 S102 之前，生成设备可以获取多个训练图像对 VGG 神经网络中各个池化层以及全连接层的参数进行调整，直到输出的结果收敛，即将训练图像作为输入，输出的视觉特征向量与关键词序列中各元素的值与预设值一致。优选地，该训练图像不仅可以包括医疗图像，还可以包括医疗图像以外其他类型的图像，例如人像图、静景图等，从而在 VGG 神经网络中，增加可识别的数量，从而提高准确率。

在 S103 中，将所述视觉特征向量以及所述关键词序列导入至预设的诊断项目识别模型，确定所述医疗图像对应的诊断项目。

在本实施例中，通过识别医疗图像中包含的关键词序列以及视觉特征向量，可以确定各个对象所对应的形状特征以及对象属性，将上述两个参数导入到预设的诊断项目识别模型，则可以确定该医疗图像所包含的诊断项目，该诊断项目具体用于表示该医疗图像所表征拍摄者的健康状况。

需要说明的是，诊断项目的个数可以基于管理员的需求进行设置，即每个医疗图像所包含的诊断项目的数量是相同的。在该情况下，管理员还可以根据不同医疗图像的图像类型，生成阈值对应的诊断项目识别模型，例如对于胸部透析图，可以采用胸部诊断项目识别模型；而 X 光膝盖透视图，则可以采用膝关节诊断项目识别模型，其中，每个识别模型所有输出结果的诊断项目的数量是固定的，即表示需要对预设的诊断项目进行识别。

在本实施例中，该诊断项目识别模型可以采用经过训练学习后的 LSTM 神经网络，在该情况下，可以将视觉特征向量以及关键词序列进行组合，构成一个医疗特征向量作为 LSTM 神经网络的输入，其中 LSTM 神经网络的层级可以与所需识别的诊断项目的个数相匹配，即每一个 LSTM 神经网络的层级对应于一个诊断项目。参见图 1c 所示，图 1c 是本申请一实施例提供的 LSTM 神经网络的结构框图，该 LSTM 神经网络中包含 N 个 LSTM 层级，每个 LSTM 层级对应 N 个诊断项目，其中 image 为基于视觉特征向量以及关键词序列生成

的医疗特征向量， $S_0 \sim S_{N-1}$ 为各个诊断项目的参数值， $p_1 \sim p_N$ 为各个参数值的正确概率，当 $\log p_i(S_{i-1})$ 收敛时，则将 S_{i-1} 所取的参数值作为该诊断项目对应的参数值，从而确定该医疗图像中各个诊断项目的值。

在 S104 中，基于诊断项目扩展模型，分别构建用于描述各个所述诊断项目的段落。

在本实施例中，生成设备在确定了各个诊断项目后，会将该诊断项目导入到诊断项目扩展模型，从而输出用于描述各个诊断项目的段落，从而患者可以通过该段落直观认知到该诊断项目的内容，提高医疗报告的可读性。

可选地，该诊断项目扩展模型可以为哈希函数，该哈希函数记录了各个诊断项目取不同参数值时对应的段落，生成设备将医疗图像对应的各个诊断项目分别导入到该哈希函数中，则可以确定该诊断项目的段落。在该情况下，生成设备只需经过哈希函数转换则可以确定段落，计算量较少，从而提高了医学报告生成的效率。

可选地，该诊断项目扩展模型可以为 LSTM 神经网络，在该情况下，生成设备会将所有诊断项目进行聚合，构成一个诊断项目向量，并将该诊断项目向量作为该 LSTM 神经网络的输入端，其中 LSTM 神经网络的层数与诊断项目的项目相同，该 LSTM 神经网络中的每一层用于输出一个诊断项目的段落，从而经过多层神经网络的输出后，则可以完成从诊断项目到段落的转化操作。通过上述方式生成段落的过程中，由于 LSTM 神经网络的输入为聚合了各个诊断项目的诊断项目向量，包含了各个诊断项目的信息，因此生成的段落能够考虑其他诊断项目的影响，从而提高了段落之间的连贯性，继而提高了整个医学报告的可读性。需要说明的是，通过 LSTM 神经网络确定段落的具体过程与 S104 相似，在此不一一赘述。

在 S105 中，根据所述段落、所述关键词序列以及所述诊断项目，生成所述医疗图像的医学报告。

在本实施例中，医学报告的生成设备在确定了该医疗图像所包含的诊断项目、描述该诊断项目的段落以及该诊断项目对应的关键词后，可以创建该医疗图像的医学报告。需要说明的是，由于诊断项目的段落已经具备了足够的可读性，可以基于诊断项目对医学报告进行模块划分，每个模块填入相应的段落，即实际用户可见的医学报告中可以只包含段落内容，而不直接体现诊断项目以及关键词。当然，生成设备可以将诊断项目、关键词以及段落进行关联显示，从而用户可以从简短精炼的关键词序列，快速确定该医学报告的具体内容，并通过诊断项目确定自身的健康状态，继而通过段落详细了解关于健康状况的评价，从不同的角度快速了解医学报告的内容，提高了医学报告的可读性以及信息获取的效率。

可选地，该医学报告可以附上医疗图像，并将关键词序列依次标记在医疗图像对应的位置，并通过标记框、列表以及分栏等方式，对照显示各个关键词对应的诊断项目以及段落信息，从而能够让用户更加直观确定该医学报告的内容。

以上可以看出，本申请实施例提供的一种医学报告的生成方法通过将医疗图像导入到预设的 VGG 神经网络，确定该医疗图像对应的视觉特征向量以及关键词序列，视觉特征向量用于表征该医疗图像中包含病症的图像特征，而关键词序列则用于确定该医疗图像中所包含的病症类型，将上述两个参数导入到诊断项目识

别模型，确定该医疗图像中所包含的诊断项目，并为每个诊断项目填充相关描述的短语以及句子，构成该诊断项目对应的段落，最后基于各个诊断项目对应的段落得到该医疗图像的医学报告。与现有的医学报告的生成方法相比，本申请实施例无需医生手动填写，可以自动根据医疗图像中包含的特征输出对应的医学报告，提高了医疗报告的生成效率，减少了人工成本，节省了患者诊疗的时间。

5 图2示出了本申请第二实施例提供的一种医学报告的生成方法S102的具体实现流程图。参见图2所示，相对于图1a述实施例，本实施例提供的一种医学报告的生成方法中S102包括S1021~S1024，具体详述如下：

在S1021中，基于所述医疗图像中各个像素点的像素值以及各个像素值的位置坐标，构建所述医疗图像的像素矩阵。

10 在本实施例中，医疗图像有多个像素点构成，每个像素点对应一个像素值，因此，基于各个像素点所在的位置坐标作为在像素矩阵的位置坐标，将像素点对应的像素值，作为像素矩阵中该像素点对应坐标的元素的值，从而可以将二维图形转换为一个像素矩阵。

需要说明的是，若该医疗图像为三基色RGB图，则可以基于医疗图像三个图层分别构建3个像素矩阵，即R图层对应一个像素矩阵，G图层对应一个像素矩阵，B图层对应一个像素矩阵，每个像素矩阵中元素的取值为0~255。当然，生成设备还可以将医疗图像进行灰度转换或二值化转换，从而将多个图层融合为一个15 图像，从而创建的像素矩阵的个数也为一个。可选地，若医疗图像为三基色RGB图，则可以将多个图层对应的像素矩阵进行融合，构成该医疗图像对应的像素矩阵，融合的方式可以为，保留三个像素矩阵中的列编号与医疗图像的横坐标一一对应，对R图层的像素矩阵的行进行扩充，每行之间填充两行空白行，并将其余两个像素矩阵的各行根据行编号的次序，依次导入扩充的各个空白行，从而构成 $3M \times N$ 的像素矩阵，其中M为医疗图像的行数，N为医疗图像的列数。

20 在S1022中，通过所述VGG神经网络的五层池化层Maxpool对所述像素矩阵进行降维操作，得到所述视觉特征向量。

在本实施例中，生成设备将构建的像素矩阵导入到VGG神经网络的五层池化层，经过五次降维操作，从而该像素矩阵所对应的视觉特征向量。需要说明的是，该池化层的卷积核可以基于像素矩阵的大小进行确定，在该情况下，生成设备记录有矩阵大小与卷积核之间的对应关系表，生成设备在构建了医疗图像对应的25 像素矩阵后，则会获取该矩阵的行数以及列数，从而确定该矩阵的尺寸，并查询该尺寸对应的卷积核尺寸，并基于该卷积核尺寸对VGG神经网络中的池化层进行调整，以使进行降维操作的过程中所使用的卷积核与像素矩阵相匹配。

在本实施例中，VGG神经网络包括用于提取视觉特征的五层池化层Maxpool以及用于确定视觉特征向量对应的关键词序列的全连接层，其中医疗图像是首先经过五层池化层后，再将降维后的向量导入到全连接层30 输出最终的关键词序列，但由于在确定诊断项目的过程中，除了需要获取描述对象以及对象属性的关键词序列外，还需要确定各个对象的视觉轮廓特征，因此生成设备会对原生的VGG神经网络进行优化，在五层池

化层后配置一个参数输出接口，以将中间变量的视觉特征向量进行导出，用于后续的操作。

在 S1023 中，将所述视觉特征向量导入所述 VGG 神经网络的全连接层，输出所述视觉特征向量对应的索引序列。

在本实施例中，生成设备会将视觉特征向量导入到 VGG 神经网络的全连接层，该全连接层中记录有各个关键词所对应的索引号，由于该 VGG 网络是经过训练学习的，因此可以通过视觉特征向量确定该医疗图像中所包含的对象以及各个对象的属性，从而通过全连接层的运算后，可以生成视觉特征向量所对应的索引序列。由于 VGG 神经网络的输出结果一般为由数字构成的向量、序列或矩阵，因此生成设备在 S1023 中并未直接输出关键词序列，而是输出关键词序列对应的索引序列，该索引序列中包含多个索引号，每个索引号对应一个关键词，从而在保证输出的结果只包含数字类型的字符的情况下，还能够确定医疗图像所对应的关键词序列。

在 S1024 中，根据关键词索引表，确定所述索引序列对应的关键词序列。

在本实施例中，生成设备存储有关键词索引表，该关键词索引表中记录了每个关键词对应的索引号，因此生成设备在确定了索引序列后，可以基于该索引序列中各个元素对应的索引号，查询与之对应的关键词，从而将索引序列转换为关键词序列。

在本申请实施例中，将五层池化层的输出作为视觉特征向量，由于经过降维操作后，可以将医疗图像中主要包含的特征通过一维的向量进行表达，从而减少了视觉特征向量的大小，提高了后续识别的效率，并输出索引序列转换为关键词序列，从而减少 VGG 模型的改造。

图 3 示出了本申请第三实施例提供的一种医学报告的生成方法 S103 的具体实现流程图。参见图 3 所示，相对于图 1a 所述实施例，本实施例提供的一种医学报告的生成方法 S103 包括 S1031~S1033，具体详述如下：

在 S1031 中，基于各个关键词在预设的语料库的序号，生成所述关键词序列对应的关键词特征向量。

在本实施例中，医学报告的生成设备存储有一记录了所有关键词的语料库，该语料库中会为每个关键词配置响应的序号，生成设备可以基于该语料库，将关键词序列转换为其对应的关键词特征向量，该关键词特征向量中包含的元素个数与关键词序列中包含的元素是一一对应的，该关键词特征向量中记录了各个关键词在语料库中对应的序号，从而可以将包含文字、英文以及数字的多种字符类型的序列转换为只包含数字类的序列，从而能够提高关键词特征序列的可运算能力。

需要说明的是，该语料库可以通过服务器下载以及用户输入的方式更新语料库中包含的关键词，对于新增的关键词，会在原有的关键词的基础上，为各个新增关键词配置相应的序号；而对于删除的关键词，则调整删除关键词序号后的所有关键词，以使整个语料库中各个关键词的序号是连续的。

在 S1032 中，分别将所述关键词特征向量以及所述视觉特征向量导入到预处理函数，得到预处理后的所述关键词特征向量以及预处理后的所述视觉特征向量；其中，所述预处理函数具体为：

$$\sigma(z_j)=\frac{e^{z_j}}{\sum_{i=1}^Me^{z_i}}$$

其中， $\sigma(z_j)$ 为所述关键词特征向量或所述视觉特征向量中第j个元素预处理后的值； z_j 为所述关键词特征向量或所述视觉特征向量中第j个元素的值； M 为所述关键词特征向量或所述视觉特征向量对应的元素个数。

在本实施例中，由于关键词序列中各个关键词在语料库中的位置差距较大时，生成的关键词特征向量中包含的序号的数值差值较大，从而不利于关键词特征向量的存储以及后续的处理，因此，在S1032中，会对关键词特征向量进行预处理，以保证关键词特征序列中所有元素的数值在预设的范围内，减少关键词特征向量的存储空间，以及减少诊断项目识别的计算量。

基于相同理由，对于视觉特征向量也可以通过预处理的方式，将视觉特征向量中各个元素的数值进行转换，以使在预设的数值范围内。

在本实施例中预处理函数的具体方式如上所述，将各个元素的值进行叠加，确定各个元素占整个向量的比例，将该比例作为该元素预处理后的参数值，从而保证了视觉特征向量以及关键词特征向量中所有元素的取值范围在0到1之间，能够减少上述两组向量的存储空间。

在S1033中，将预处理后的所述关键词特征向量以及预处理后的所述视觉特征向量作为所述诊断项目识别模型的输入，输出所述诊断项目。

在本实施例中，生成设备将预处理后的关键词向量以及预处理后的视觉特征向量作为诊断项目识别模型的输入，由于经过上述处理后，上述两组向量的值在预设的范围内，从而减少了每个元素所需分配的字节数，有效对整个向量的大小进行控制，在诊断项目识别模型进行计算时，也能够减少无效的位数的读取操作，提高了处理的效率，而上述向量中每个元素的参数值并未发生本质的变化，而是等比例地缩小，依然可以确定诊断项目。

需要说明的是，上述诊断项目的识别模型可以参数LSTM神经网络以及上述各实施例中所提供的神经网络，具体实现过程可参见上述实施例，在此不再一一赘述。

在本申请实施例中，通过对关键词序列以及视觉特征向量进行预处理，从而提高了医学报告的生成效率。

图4示出了本申请第四实施例提供的一种医学报告的生成方法的具体实现流程图。参见图4所示，相对于图1a~图3所述实施例，本实施例提供的一种医学报告的生成方法中还包括：S401~S403，具体详述如下：

进一步地，在所述将所述视觉特征向量以及所述关键词序列导入至预设的诊断项目识别模型，确定所述医疗图像对应的诊断项目之前，还包括：

在S401中，获取多个训练图像的训练视觉向量、训练关键词序列以及训练诊断项目。

在本实施例中，医学报告的生成设备会获取多个预设的训练图像的训练视觉向量、训练关键词序列以及训练诊断项目。优选地，该训练图像的个数应大于1000个，从而提高该LSTM神经网络的识别准确性。需

要强调的是，该训练图像可以为历史医疗图像，还可以为不限于医疗类型的其他图像，从而提高了 LSTM 神经网络的可识别对象的种类数。

需要说明的是，各个训练图像的训练诊断项目的格式是相同的，即训练诊断项目的项数是相同的。若任一训练图像由于拍摄角度的问题无法解析出部分训练诊断项目，则该训练诊断项目的值为空，从而保证了在
5 对 LSTM 神经网络进行训练时，各个通道输出的参数的含义是固定的，提高了 LSTM 神经网络的准确性。

在 S402 中，将所述训练视觉向量以及所述训练关键词序列作为长短期 LSTM 神经网络的输入，将所述训练诊断项目作为所述 LSTM 神经网络的输出，对所述 LSTM 神经网络内的各个学习参数进行调整，以使所述 LSTM 神经网络满足收敛条件；所述收敛条件为：

$$\theta^* = \arg \max_{\theta} \sum_{Stc} \log p(Visual, Keyword|Stc; \theta)$$

其中， θ^* 为调整后的所述学习参数；*Visual*为所述训练视觉向量；*Keyword*为所述训练关键词序列；*Stc*为
10 所述训练诊断项目； $p(Visual, Keyword|Stc; \theta)$ 为当所述学习参数的值为 θ 时，将所述所述训练视觉向量以及所述训练关键词序列导入到所述 LSTM 神经网络，输出结果为该所述训练诊断项目的概率值； $\arg \max_{\theta} \sum_{Stc} \log p(Visual, Keyword|Stc; \theta)$ 为所述概率值取最大值时所述学习参数的取值。

在本实施例中，LSTM 神经网络中包含多个神经层，每个神经层设置有相应的学习参数，通过调整学习参数的参数值能够适应不同输入类型以及输出类型。当学习参数设置为某一参数值时，将多个训练对象的对象图像输入到该 LSTM 神经网络，将对应输出一各个对象的对象属性，生成设备会将输出的诊断项目与训练
15 诊断项目进行比对，确定本次输出是否正确，并且基于多个训练对象的输出结果，得到该学习参数取该参数值时输出结果正确的概率值。生成设备会调整该学习参数，以使该概率值取最大值，则表示该 LSTM 神经网络已经调整完毕。

在 S403 中，将调整后的 LSTM 神经网络作为诊断项目识别模型。

20 在本实施例中，终端设备将调整了学习参数后的 LSTM 神经网络作为诊断项目识别模型，提高了诊断项目识别模型识别的准确率。

在本申请实施例中，通过训练对象对 LSTM 神经网络进行训练，选取输出结果正确的概率值最大时对应的学习参数作为 LSTM 神经网络中学习参数的参数值，从而提高了诊断项目识别的准确性，进一步医学报告的准确率。

25 图 5 示出了本申请第五实施例提供的一种医学报告的生成方法的具体实现流程图。参见图 5 所示，相对于图 1a 所述实施例，本实施例提供的一种医学报告的生成方法包括：S501~S50，具体详述如下：

在 S501 中，接收待识别的医疗图像。

由于 S501 与 S101 的实现方式相同，具体参数可参见 S101 的相关描述，在此不再赘述。

在 S502 中，对所述医疗图像进行二值化处理，得到二值化后的医疗图像。

在本实施例中，生成设备会对医疗图像进行二值化处理，以使医疗图像中各个对象的边缘更加明显，从而方便确定各个对象的轮廓，以及每个对象的内部结构，方便实现视觉特征向量以及关键词序列的提取操作。

在本实施例中，二值化的阈值可以根据用户的需求进行设置，生成设备也可以通过确定该医疗图像的类型和/或医疗图像中各个像素点的平均像素值，确定该二值化的阈值，从而提高了二值化后医疗图像的显示效果。

在 S503 中，识别二值化后的所述医疗图像的边界，将所述医疗图像划分为多个医疗子图像。

在本实施例中，生成设备可以通过预设的边界识别算法，从二值化后的医疗图像中提取各个对象的边界，从而基于识别得到的边界对医疗图像进行划分，从而得到每个对象独立的医疗子图像。当然，若某几个对象是相互关联的，且边界是重叠或相邻的，则上述几个对象可以集成在一个医疗子图像内。通过对不同对象进行区域划分，从而在对某一对象进行视觉特征以及关键词提取的操作中，减少其他对象对其的影响。

进一步地，所述将所述医疗图像导入预设的 VGG 神经网络，得到所述医疗图像的视觉特征向量以及关键词序列，包括：

在 S504 中，将各个医疗子图像分别导入所述 VGG 神经网络，得到所述医疗子图像的视觉特征分量以及关键词子序列。

在本实施例中，生成设备会将基于医疗图像分割得到的各个医疗子图像分别导入 VGG 神经网络，从而分别得到各个医疗子图像对应的视觉特征分量以及关键词子序列，该视觉特征分量用于表征该医疗子图像中对象的形状、轮廓特征，而关键词子序列则用于表示该医疗子图像中包含的对象。通过将医疗图像划分，分别导入到 VGG 神经网络内，能够减少每次 VGG 神经网络运算的数据量，从而大大减少了处理时间，提高输出效率。并且由于基于边界进行划分，可以有效地删除大部分无效的背景区域图像，从而整体的数据处理量会大幅减少。

在 S505 中，基于各个所述视觉特征分量生成所述视觉特征向量，以及基于各个所述关键词子序列构成所述关键词序列。

在本实施例中，将各个医疗子图像的视觉特征分量进行合并，构成该医疗图像的视觉特征向量；同样的，将各个医疗子图像的关键词子序列进行合并，构成该医疗图像的关键词序列。需要说明的是，在合并的过程中，某一医疗子图像的视觉特征分量在合并后的视觉特征向量中的位置与该医疗子图像的关键词子序列在合并后的关键词序列中的位置是对应的，从而保持两者之间的关联关系。

在 S506 中，将所述视觉特征向量以及所述关键词序列导入至预设的诊断项目识别模型，确定所述医疗图像对应的诊断项目。

在 S507 中，基于诊断项目扩展模型，分别构建用于描述各个所述诊断项目的段落。

在 S508 中，根据所述段落、所述关键词序列以及所述诊断项目，生成所述医疗图像的医学报告。

由于 S506~S508 与 S103~S105 的实现方式相同，具体参数可参见 S103~S105 的相关描述，在此不再赘述。

在本申请实施例中，通过对医疗图像进行边界划分，得到多个医疗子图像并分别确定各个医疗子图像对应的视觉特征分类以及关键词子序列，最后构建得到医疗图像的视觉特征向量以及关键词序列，从而减少了 VGG 神经网络的数据处理量，提高了生成效率。

应理解，上述实施例中各步骤的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能性和内在逻辑确定，而不应对本申请实施例的实施过程构成任何限定。

图 6 示出了本申请一实施例提供的一种医学报告的生成设备的结构框图，该医学报告的生成设备包括的各单元用于执行图 1a 对应的实施例中的各步骤。具体请参阅图 1a 与图 1a 所对应的实施例中的相关描述。为了便于说明，仅示出了与本实施例相关的部分。

参见图 6，所述医学报告的生成设备包括：

10 医疗图像接收单元 61，用于接收待识别的医疗图像；

特征向量获取单元 62，用于将所述医疗图像导入预设的视觉几何组 VGG 神经网络，得到所述医疗图像的视觉特征向量以及关键词序列；

诊断项目识别单元 63，用于将所述视觉特征向量以及所述关键词序列导入至预设的诊断项目识别模型，确定所述医疗图像对应的诊断项目；

15 描述段落确定单元 64，用于基于诊断项目扩展模型，分别构建用于描述各个所述诊断项目的段落；

医学报告生成单元 65，用于根据所述段落、所述关键词序列以及所述诊断项目，生成所述医疗图像的医学报告。

可选地，所述特征向量获取单元 62 包括：

20 像素矩阵构建单元，用于基于所述医疗图像中各个像素点的像素值以及各个像素值的位置坐标，构建所述医疗图像的像素矩阵；

视觉特征向量生成单元，用于通过所述 VGG 神经网络的五层池化层 Maxpool 对所述像素矩阵进行降维操作，得到所述视觉特征向量；

索引序列生成单元，用于将所述视觉特征向量导入所述 VGG 神经网络的全连接层，输出所述视觉特征向量对应的索引序列；

25 关键词序列生成单元，用于根据关键词索引表，确定所述索引序列对应的关键词序列。

可选地，所述诊断项目识别单元 63 包括：

关键词特征向量构建单元，用于基于各个关键词在预设的语料库的序号，生成所述关键词序列对应的关键词特征向量；

30 预处理单元，用于分别将所述关键词特征向量以及所述视觉特征向量导入到预处理函数，得到预处理后的所述关键词特征向量以及预处理后的所述视觉特征向量；其中，所述预处理函数具体为：

$$\sigma(z_j) = \frac{e^{z_j}}{\sum_{i=1}^M e^{z_i}}$$

其中, $\sigma(z_j)$ 为所述关键词特征向量或所述视觉特征向量中第 j 个元素预处理后的值; z_j 为所述关键词特征向量或所述视觉特征向量中第 j 个元素的值; M 为所述关键词特征向量或所述视觉特征向量对应的元素个数;

预处理向量导入单元, 用于将预处理后的所述关键词特征向量以及预处理后的所述视觉特征向量作为所述诊断项目识别模型的输入, 输出所述诊断项目。

可选地, 所述医学报告的生成设备还包括:

训练参数获取单元, 用于获取多个训练图像的训练视觉向量、训练关键词序列以及训练诊断项目;

学习参数训练单元, 用于将所述训练视觉向量以及所述训练关键词序列作为长短期 LSTM 神经网络的输入, 将所述训练诊断项目作为所述 LSTM 神经网络的输出, 对所述 LSTM 神经网络内的各个学习参数进行调整, 以使所述 LSTM 神经网络满足收敛条件; 所述收敛条件为:

$$\theta^* = \arg \max_{\theta} \sum_{Stc} \log p(Visual, Keyword | Stc; \theta)$$

其中, θ^* 为调整后的所述学习参数; $Visual$ 为所述训练视觉向量; $Keyword$ 为所述训练关键词序列; Stc 为所述训练诊断项目; $p(Visual, Keyword | Stc; \theta)$ 为当所述学习参数的值为 θ 时, 将所述训练视觉向量以及所述训练关键词序列导入到所述 LSTM 神经网络, 输出结果为该所述训练诊断项目的概率值; $\arg \max_{\theta} \sum_{Stc} \log p(Visual, Keyword | Stc; \theta)$ 为所述概率值取最大值时所述学习参数的取值;

诊断项目识别模型生成单元, 用于将调整后的 LSTM 神经网络作为诊断项目识别模型。

可选地, 所述医学报告的生成设备还包括:

二值化处理单元, 用于对所述医疗图像进行二值化处理, 得到二值化后的医疗图像;

边界划分单元, 用于识别二值化后的所述医疗图像的边界, 将所述医疗图像划分为多个医疗子图像;

所述特征向量获取单元 62 包括:

医疗子图像识别单元, 用于将各个医疗子图像分别导入所述 VGG 神经网络, 得到所述医疗子图像的视觉特征分量以及关键词子序列;

特征向量合成单元, 用于基于各个所述视觉特征分量生成所述视觉特征向量, 以及基于各个所述关键词子序列构成所述关键词序列。

因此, 本申请实施例提供的医学报告的生成设备同样无需医生手动填写, 可以自动根据医疗图像中包含的特征输出对应的医学报告, 提高了医疗报告的生成效率, 减少了人工成本, 节省了患者诊疗的时间。

图 7 是本申请另一实施例提供的一种医学报告的生成设备的示意图。如图 7 所示, 该实施例的医学报告的生成设备 7 包括: 处理器 70、存储器 71 以及存储在所述存储器 71 中并可在所述处理器 70 上运行的计算机可读指令 72, 例如医学报告的生成程序。所述处理器 70 执行所述计算机可读指令 72 时实现上述各个医

学报告的生成方法实施例中的步骤，例如图 1a 所示的 S101 至 S105。或者，所述处理器 70 执行所述计算机可读指令 72 时实现上述各装置实施例各单元的功能，例如图 6 所示模块 61 至 65 功能。

示例性的，所述计算机可读指令 72 可以被分割成一个或多个单元，所述一个或者多个单元被存储在所述存储器 71 中，并由所述处理器 70 执行，以完成本申请。所述一个或多个单元可以是能够完成特定功能的一系列计算机可读指令指令段，该指令段用于描述所述计算机可读指令 72 在所述医学报告的生成设备 7 中的执行过程。例如，所述计算机可读指令 72 可以被分割成医疗图像接收单元、特征向量获取单元、诊断项目识别单元、描述段落确定单元以及医学报告生成单元，各单元具体功能如上所述。

所述医学报告的生成设备 7 可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。所述医学报告的生成设备可包括，但不仅限于，处理器 70、存储器 71。本领域技术人员可以理解，图 7 仅仅是医学报告的生成设备 7 的示例，并不构成对医学报告的生成设备 7 的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件，例如所述医学报告的生成设备还可以包括输入输出设备、网络接入设备、总线等。

所称处理器 70 可以是中央处理单元（Central Processing Unit，CPU），还可以是其他通用处理器、数字信号处理器（Digital Signal Processor，DSP）、专用集成电路（Application Specific Integrated Circuit，ASIC）、现场可编程门阵列（Field-Programmable Gate Array，FPGA）或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

所述存储器 71 可以是所述医学报告的生成设备 7 的内部存储单元，例如医学报告的生成设备 7 的硬盘或内存。所述存储器 71 也可以是所述医学报告的生成设备 7 的外部存储设备，例如所述医学报告的生成设备 7 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 71 还可以既包括所述医学报告的生成设备 7 的内部存储单元也包括外部存储设备。所述存储器 71 用于存储所述计算机可读指令以及所述医学报告的生成设备所需的程序和数据。所述存储器 71 还可以用于暂时地存储已经输出或者将要输出的数据。

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。

以上所述实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围，均应包含在本申请的保护范围之内。