



(12) 发明专利申请

(10) 申请公布号 CN 115619706 A

(43) 申请公布日 2023.01.17

(21) 申请号 202210742157.6

G06N 3/04 (2006.01)

(22) 申请日 2022.06.27

G06N 3/08 (2006.01)

(71) 申请人 青岛科技大学

地址 266000 山东省青岛市崂山区松岭路  
99号

(72) 发明人 许灿辉 张洪红 李玉腾 史操  
程远志 刘宇博 杨琦

(74) 专利代理机构 青岛汇智海纳知识产权代理  
有限公司 37335

专利代理师 万桂斌

(51) Int. Cl.

G06T 7/00 (2017.01)

G06T 7/11 (2017.01)

G06T 7/13 (2017.01)

G06T 7/187 (2017.01)

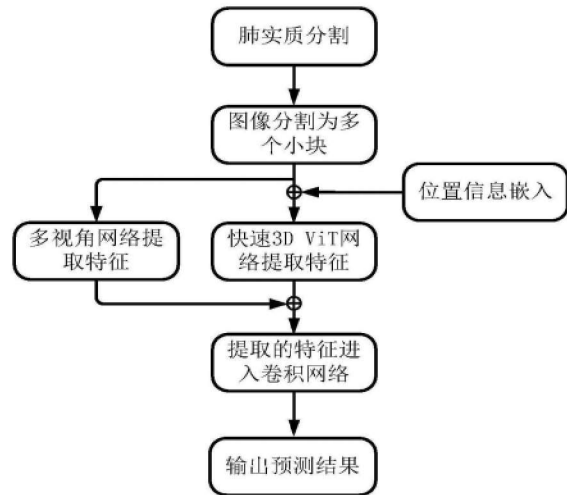
权利要求书1页 说明书8页 附图3页

(54) 发明名称

一种基于深度学习的肺结节检测方法

(57) 摘要

本发明提出一种基于深度学习的肺结节检测方法,包括:A、对三维肺部CT图像进行实质分割;B、将肺实质图像分块之后序列化,最终得到稀疏化矩阵;C、提取稀疏矩阵的三视角信息用于假阳性剔除;D、搭建网络模型;E、使用训练好的网络模型进行肺结节的检测并输出检测结果。本发明结合三视图去假阳辅助模块,直接嵌入到端对端的框架中,利用稀疏化后的三维图像三视图自注意力信息帮助假阳性区域的剔除。一方面,解决了复杂的3D场景中学习注意力困难的问题。另一方面,将检测和假阳性区域剔除设计成一个可以端对端训练的模型,提高了检测速度和检测精度,降低了模型整体的复杂度,在训练模型时可以统一损失函数训练,提高了模型训练收敛的速度。



1. 一种基于深度学习的肺结节检测方法,其特征在于包括:

步骤A、对三维肺部CT图像进行实质分割,去掉无关组织得到肺实质图像;

步骤B、将肺实质图像分块之后序列化,输入到快速3D ViT网络中进行提取特征,输出多尺度特征图;同时去掉三维肺部实质分块图像中梯度为0及灰度变化不大的像素区域,得到稀疏化矩阵;

步骤C、提取稀疏矩阵的三视角信息用于假阳性筛除;

步骤D、搭建网络模型并训练网络参数;

步骤E、使用训练好的网络模型进行肺结节的检测并输出检测结果。

2. 根据权利要求1所述的基于深度学习的肺结节检测方法,其特征在于,所述步骤B中,多尺度特征图的构建过程如下:利用快速3D ViT作为骨干网络输出的最后一层特征信息分别构建4种尺度的特征图,这4种尺度的特征图之间并没有连接路径,每一层都由骨干网络的最后一层输出直接构建,尺寸逐渐扩大为上一层的两倍,第二层的尺寸与骨干网络最后一层的输出一致。

3. 根据权利要求1所述的基于深度学习的肺结节检测方法,其特征在于,所述步骤B中序列化过程如下:将分割后的图像向量做线性变换,将三维图像投影到二维空间再展平到一维,并嵌入一维位置信息。

4. 根据权利要求1所述的基于深度的图像向量做线性变换,将三维图像投影到二维空间再展平到一维,并嵌入一维位置信息。学习的肺结节检测方法,其特征在于,所述步骤B中,稀疏化矩阵生成过程如下:

步骤B1、对三维肺部CT图像经行处理;

步骤B2、初始化水平集方程;

步骤B3、根据初始化的水平集方程计算前景和背景灰度均值,用前面计算的前景和背景均值对水平集每一个点进行调整;

步骤B4、更新水平集方程,并根据水平集方程画出初始轮廓。

5. 根据权利要求1所述的基于深度学习的肺结节检测方法,其特征在于,所述步骤A中分割方法如下:

步骤A1、将三维肺部CT图像中的无关区域进行裁剪,仅保留肺实质图像所在区域;

步骤A2、对裁剪后的图像进行二值化处理;

步骤A3、消除肺部的空孔洞,提取肺实质的最大连通区域,并对图像进行二值化反转处理获得肺实质的掩膜图像;

步骤A4、将掩膜图像与步骤A1获得的图象相乘,得到完整的肺实质图像。

6. 根据权利要求1所述的基于深度学习的肺结节检测方法,其特征在于,所述步骤B中,3D ViT网络的编码层由归一化层、多头注意力层,再次归一化层,多层感知机层组成。

7. 根据权利要求1所述的基于深度学习的肺结节检测方法,其特征在于,所述步骤C中三视角为正视、俯视及侧视。

## 一种基于深度学习的肺结节检测方法

### 技术领域

[0001] 本发明属于医学图像处理技术领域,具体地涉及一种肺部结节检测方法。

### 背景技术

[0002] 肺癌是世界死亡率较高的疾病之一。肺结节的检测对肺癌的预测有重大意义,但由于肺结节形态各异,和正常组织之间的差别并不明显,即便对于医生来说也需要详细地分辨才能得出具体的结论。随着人工智能技术的广泛应用,利用深度学习的方法检测肺部结节成为了一种快速检测手段,通过在数据集上对网络模型的训练,对肺结节实现自动检测识别,可以提高效率,帮助医生减轻负担。

[0003] 肺结节检测是一个比较困难的三维物体检测问题,并且,有限的GPU内存也是三维物体检测面临的技术难题之一。同时,3D数据的标注通常比二维数据的标注困难的多,这会导致深度学习模型产生过拟合的问题。另外,肺部CT图像通常包含较多的外部无关组织,比如患者的躯干,以及肺部器官内部包含血管、气管等组织,这些都很容易被误识别为肺结节,造成假阳性较多的现象。因此在使用深度学习检测方法得出的预测结果中存在较多的假阳性结果,导致检测的准确率降低。

### 发明内容

[0004] 本发明为解决现有肺部结节检测方法存在较多的假阳性结果,导致检测的准确率降低的问题,提出一种基于深度学习的肺结节检测方法,其采用如下技术方案予以实现:

[0005] 一种基于深度学习的肺结节检测方法,包括:

[0006] 步骤A、对三维肺部CT图像进行实质分割,去掉无关组织得到肺实质图像;

[0007] 步骤B、将肺实质图像分块之后序列化,输入到快速3D ViT网络中进行提取特征,输出多尺度特征图;同时去掉三维肺部实质分块图像中梯度为0及灰度变化不大的像素区域(灰度值变化范围在0~255,本发明将灰度值变化在0~50的范围定义为灰度变化不大的像素区域),得到稀疏化矩阵;

[0008] 步骤C、提取稀疏矩阵的三视角信息用于假阳性筛除;

[0009] 步骤D、搭建网络模型并训练网络参数;

[0010] 步骤E、使用训练好的网络模型进行肺结节的检测并输出检测结果。

[0011] 进一步地,所述步骤B中,多尺度特征图的构建过程如下:利用快速3D ViT作为骨干网络输出的最后一层特征信息分别构建4种尺度的特征图,这4种尺度的特征图之间并没有连接路径,每一层都由骨干网络的最后一层输出直接构建,尺寸逐渐扩大为上一层的两倍,第二层的尺寸与骨干网络最后一层的输出一致。

[0012] 进一步地,所述步骤B中序列化过程如下:将分割后的图像向量做线性变换,将三维图像投影到二维空间再展平到一维,并嵌入一维位置信息。

[0013] 进一步地,所述步骤B中,稀疏化矩阵生成过程如下:

[0014] 步骤B1、对三维肺部CT图像经行处理;

- [0015] 步骤B2、初始化水平集方程；
- [0016] 步骤B3、根据初始化的水平集方程计算前景和背景灰度均值，用前面计算的前景和背景均值对水平集每一个点进行调整；
- [0017] 步骤B4、更新水平集方程，并根据水平集方程画出初始轮廓。
- [0018] 进一步地，所述步骤A中分割方法如下：
- [0019] 步骤A1、将三维肺部CT图像中的无关区域进行裁剪，仅保留肺实质图像所在区域；
- [0020] 步骤A2、对裁剪后的图像进行二值化处理；
- [0021] 步骤A3、消除肺部的空孔洞，提取肺实质的最大连通区域，并对图像进行二值化反转处理获得肺实质的掩膜图像；
- [0022] 步骤A4、将掩膜图像与步骤A1获得的图像相乘，得到完整的肺实质图像。
- [0023] 进一步地，所述步骤B中，3D ViT网络的编码层由归一化层、多头注意力层，再次归一化层，多层感知机层组成。
- [0024] 进一步地，所述步骤C中三视角为正视、俯视及侧视。
- [0025] 与现有技术相比，本发明的优点和积极效果在于：
- [0026] 1、考虑到三维图像的数据量比较庞大，本发明通过快速3D ViT网络结构提取特征图像信息生成多尺度特征图，只采用骨干网络最后一层的输出来分别构建四层不同尺度的特征图，如此，在计算过程中不改变每一层特征输出的尺寸，将语义特征集成到最后一层输出上，仅利用最后一层输出特征来构建多尺度特征图，输出速度大幅提高。
- [0027] 2、结合三视图去假阳辅助模块，直接嵌入到端对端的框架中，利用稀疏化后的三维图像三视图自注意力信息帮助假阳性区域的筛除。一方面，解决了复杂的3D场景中学习注意力困难的问题。另一方面，本发明采用的是在检测框架中嵌入包含可以用于假阳性筛除信息的模块，将检测和假阳性区域筛除设计成一个可以端对端训练的模型，提高了检测速度和检测精度，降低了模型整体的复杂度，在训练模型时可以统一损失函数训练，提高了模型训练收敛的速度。

## 附图说明

- [0028] 图1是本发明采用深度学习检测肺结节的方法示意性流程图；
- [0029] 图2是本发明肺实质分割步骤图；
- [0030] 图3是本发明采用深度学习检测肺结节的网络结构图；
- [0031] 图4是本发明水平集图像稀疏化处理过程
- [0032] 图5是本发明三视角自注意力算法过程
- [0033] 图6是本发明肺结节检测结果示例
- [0034] 图7是假阳性区域筛除前后对比，其中(a)图为没有嵌入假阳性模块的结果，灰色框内的为假阳性结节，(b)图为嵌入假阳性模块的结果。

## 具体实施方式

- [0035] 本发明提出了一种基于深度学习的肺结节检测方法。首先，将肺部CT图像进行肺部实质分割处理，去掉无关组织仅保留肺实质区域；然后将肺实质图像分块之后序列化，嵌入可学习的一维位置信息后输入到快速3D ViT中提取特征，输出多尺度特征图；同时去掉

3D肺部实质分块图像中大部分梯度为0即灰度变化不大的地方,因为这些位置通常没有病变;去掉大部分背景后的三维图像变为稀疏化的矩阵。提取稀疏矩阵的三视角信息用于假阳性剔除;接着,训练网络参数并使用训练好的网络模型进行肺结节的检测;最终,输出预测结果,即候选区域所在位置坐标、半径以及是否为肺结节的概率。

[0036] 为了便于理解本发明,下文结合附图图1实施例对本发明做进一步的描述,但本发明的保护范围不限于以下具体实施例。

[0037] 一、参考图1,本发明的主要实施步骤如下:

[0038] 第一步、基于统计的肺部CT图像实质分割。

[0039] 肺部CT图像通常包含躯干、肺实质、肺气管、肺血管。肺实质在图像上是面积较大的低灰度连通区域,而其他无关结构区域的灰度值较高。肺部的无关组织也可能呈球形,造成对肺结节检测的干扰,本实施例对CT图像先进行肺实质分割处理。

[0040] 肺部CT图像原始的数据集定义为:

$$[0041] \quad \text{CTSet} = \{S_i \in \mathbb{R}^{H \times W \times C}, i = 1 \dots N\} \quad (1)$$

[0042] 其中, $S_i$ 表示第*i*个CT图像样本(Sample),数据集中的每一个样本都是三维图像, $H \times W$ 为空间分辨率, $C$ 为通道数,即:CT切片数。 $\mathbb{R}$ 表示三维向量空间。每个样本 $S_i$ 都由多张切片组成,即:

$$[0043] \quad S = \{\text{slice}_j \in \mathbb{R}^{H \times W}, j = 1 \dots C\} \quad (2)$$

[0044] CT扫描设备(Computed Tomography)的一致性,以及成像过程中医生对于每位患者检查姿势调整的统一指示,因此肺部CT成像的布局是规律的,肺部整体位于图像中心位置,四周为其余无关组织。如图2所示,肺实质区域位于当前样本切片的中心位置,肺实质区域中心坐标变量 $(x_c, y_c)$ 服从以下分布:

$$[0045] \quad x_c, y_c \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) \quad (3)$$

[0046]  $(x_c, y_c)$ 服从的是二维高斯分布或者二维正态分布,五个参数,从左到右依次为 $x_c$ 的期望, $y_c$ 的期望, $x_c$ 的方差, $y_c$ 的方差, $x_c$ 和 $y_c$ 的相关系数)每张切片的中心坐标变量都服从同一分布,并且互相独立,即它们是独立同分布(Independently identical distribution)。同时,宽高变量 $W, H$ 分别服从:

$$[0047] \quad H \sim N(\mu_3, \sigma_3)$$

$$[0048] \quad W \sim N(\mu_4, \sigma_4) \quad (4)$$

[0049] 它们之间同样是独立同分布。

[0050] 基于以上,结合图2,肺实质分割的步骤如下:

[0051] (a)到(b):基于宽高变量 $W, H$ 服从的分布,确定估计值 $w, h$ ,将原始CT图像中的无关区域进行裁剪,仅保留肺实质所在区域。

[0052] (b)到(c):裁剪后的图像进行二值化处理。通过适当的阈值选取,使得整个图像呈现出明显的黑白效果,凸显出肺实质区域的轮廓。

[0053] (c)到(d):消除肺部的空孔洞,提取肺实质的最大连通区域,并对图像进行二值化反转处理获得肺实质的掩膜图像。

[0054] (d)到(e):将掩膜图像与图像(b)相乘,得到完整的肺实质图像。

[0055] 预处理算法如下:

[0056] <1>输入完整的原始肺部CT图像 $a, a = \text{image}(\text{original})$

[0057] original指的是原始的肺部CT图像,image()为输入函数,a是检测模型的输入

[0058] <2>根据肺实质区域与其他无关区域灰度值的差异对图像a进行裁剪,仅保留肺实质所在区域,得到图像b, $b=crop(a, (h,w))$

[0059] crop()是裁剪函数,a是输入的图像,h,w分别是裁剪之后图像的高,宽尺寸,b是裁剪之后得到的图像

[0060] <3>选取合适的阈值T,对图像b进行二值化处理。b中灰度值大于等于T的像素划分为前景,标记为黑色;灰度值小于T的像素为背景,标记为白色。得到图像c, $c=binarize(b, T)$ 。

[0061] binarize()是二值化函数,T为二值化的灰度阈值,c是b二值化之后得到的图像

[0062] <4>在图像(c)中,依旧包含很多孔洞。通过连通计算将像素点个数最大的连通区域标记,提取最大连通区域,消除孔洞。然后对二值化图像取反,得到肺实质的掩膜(d), $d=maxfill(c)$ 。

[0063] maxfill()代表提取最大连通区域的过程,抽象为函数,d是消除孔洞之后的图像

[0064] <5>肺实质掩膜(d)与图像(b)相乘输出分割后的肺实质图像(e), $e=multiply(d, b)$ 。

[0065] <6>算法结束。

[0066] 第二步、肺实质图像分块处理。

[0067] 考虑到Transformer结构不能直接处理三维图像,在训练过程中对内存的要求很高。本实施例对输入的图像进行分块处理,具体如下:

[0068] 经过实质分割得到的肺实质图像数据集定义为:

$$[0069] \quad SCTSet = \{Seg_i \in R^{h \times w \times c}, i = 1 \dots N\} \quad (5)$$

[0070] 图像分块的尺寸大小为 $p \times p \times c$ ,即分割后得到的patches的尺寸为 $p \times p \times c$ 。

[0071] 肺实质图像的尺寸为 $Seg \in R^{h \times w \times c}$ ,将其分成 $p \times p \times c$ 的patches,所以实际上会有N个patches,其中:

$$[0072] \quad N = (\lfloor h/p \rfloor + 1) \times (\lfloor w/p \rfloor + 1) \times (\lfloor C/c \rfloor + 1) \quad (6)$$

[0073] 分割后得到的全部patches的维度为 $N \times p \times p \times c$ 。然后将三维的patches展平,即:

$$[0074] \quad P: R^{p \times p \times c} \rightarrow R^{p^2 \times c} \quad (7)$$

[0075] 接下来需要将得到的patches进行一个线性嵌入:即高维转换为低维的过程。对每一个patch向量做一个线性变换,即全连接层,降维后的维度为D。

$$[0076] \quad X_p = \text{linear-embedding}(P, (p^2 \times c, D)) \quad (8)$$

[0077] 其中 $X_p \in R^{N \times D}$ 。

[0078] 为了保证输入图像patches之间的空间位置信息,还需要对图像块嵌入中添加一个位置编码向量,

$$[0079] \quad Z = \left\{ X_p^1 + E_{position}^1, X_p^2 + E_{position}^2, X_p^3 + E_{position}^3, \dots, X_p^i + E_{position}^i \right\}, i = N \quad (9)$$

[0080] 其中, $E_{position}^i \in R^{N \times D}$ 。

[0081] 得到的 $Z \in R^{N \times D}$ 就是要下一步要输入到网络模型中的向量序列。

[0082] 第三步、搭建网络模型。

[0083] 网络模型主要由特征提取,稀疏化算法,自注意力模块,卷积网络部分组成。

[0084] 特征提取的骨干网络为快速3D ViT网络。将预处理的特征 $Z \in \mathbb{R}^{N \times D}$ 输入到编码层中进行处理,每一个编码层主要由归一化(Norm)、多头注意力(Multi-Head Attention),再次归一化(Norm),多层感知机(MLP)四个计算工序组成。在特征提取过程中,依次叠加L层这样的编码结构。特征提取的过程表示为:

$$[0085] \quad Z_1 = \text{Encoder}(Z_{1-1}), 1=1, 2, 3 \dots L \quad (10)$$

[0086] 最后 $Z_L$ 输出的特征将其定义为 $F_1$ 且 $F_1 \in \mathbb{R}^{N \times D}$ ,由特征 $F_1$ 来创建特征金字塔。

[0087] 如图3中,为了更加全面的检测所有疑似结节区域,通过卷积或者反卷积的方式对特征 $F_1$ 进行解码,获得多尺度信息,生成四种不同尺度的特征输出,用于候选检测,确保能检测到各种尺寸不一的结节。

[0088]  $F_1$ 分别生成 $L_1, L_2, L_3, L_4$ 四层特征图,每一层都由 $F_1$ 生成,层与层之间没有连接的路径,减少了生成特征图时的计算量。从 $L_1$ 到 $L_4$ 的尺寸依次减少为上一层的一半且这四个尺度的特征输出是平行的,且彼此之间不存在自上而下或者自下而上的相连的路径。

[0089] 对上一步得到的P(肺实质图像分块后得到的patches)同时进行水平集算法处理得到包含肺结节或者其他孔洞组织的稀疏矩阵。对于肺结节检测而言,肺实质图像包含的背景区域较多,给肺结节的检测带来较大干扰,本发明通过水平集算法将图像稀疏化处理。在整个三维切片内部将所有疑似肺结节区域划分出来,原先的P就变为一个稀疏矩阵。

[0090] 为了得到更多的特征信息对假阳性区域进行筛除,将P的三种视图(正视图、侧视图、俯视图)分别输入网络,从多个视角进行联合学习肺结节的特征。

[0091] 本发明利用三视图进行自注意力计算构建了用以假阳性区域筛除的模块,可对假阳性区域进行充分的判别。如图5所示,A,S,C分别代表三维图像的俯视图,侧视图,正视图,分别通过 $3 \times 3$ 的卷积,得到Q,K,V。为了得到Q,K之间的相关性联系,将二者展平之后做乘积并进行softmax函数的计算,得到注意力分数,然后将注意力分数乘到展平之后的输出上,得到输出特征 $F_2$ 。

[0092] 如图5所示,三个视角的信息分别输入到自注意力机制模块中,以正视角的特征信息为主,俯视图和侧视图的特征信息为辅计算注意力分数,输出特征 $F_2$ ,然后融合到多尺度特征图中。

[0093] 特征 $F_2$ 和 $F_1$ 输出的多尺度特征 $L_1 \sim L_4$ 进行concat融合。

[0094] 根据 $L_1 \sim L_4$ 输出特征的尺寸,需要将特征 $F_2$ 的尺寸进行调整,使得 $F_2$ 的尺寸能够分别于 $L_1 \sim L_4$ 融合。这里使用卷积运算来调整 $F_2$ 的尺寸:

$$[0095] \quad F_2^i = \text{reshape}(F_2, L_i) \quad (11)$$

$$[0096] \quad L_i \oplus F_2^i : \text{concat}(L_i, F_2^i), i=1, 2, 3, 4 \quad (12)$$

[0097] 融合之后得到新的特征,这里的特征既包含了经过编码层提取的特征,又利用了三个视角之间的特征对下一步的检测进一步的选择。

[0098] 本发明采用的是在检测框架中嵌入包含可以用于假阳性筛除信息的模块,将检测和假阳性区域筛除设计成一个可以端对端训练的模型,极大提高了检测速度和检测精度,降低了模型整体的复杂度,在训练模型时可以统一损失函数训练,提高了模型训练收敛的

速度。)

[0099] 随后经过两层卷积网络运算之后,输出四维的tensor,为预测的proposals,最后两个维度分别对应于锚框和回归量,五个回归值分别为 $(x, y, z, r, p)$ 。 $(x, y, z)$ 分别代表坐标位置,  $r$ 为该结节的半径,  $p$ 为该区域是结节的概率。

[0100] 将 $L_1 \sim L_4$ 分别输入到两层卷积组成的网络,然后连接全连接层,最终用softmax函数输出结果:

$$[0101] \quad f_i = \text{ConvNet}(L_i) \quad (13)$$

$$[0102] \quad V_i = \text{FC}(f_i), i = 1, 2, 3, 4 \quad (14)$$

[0103]  $V_i$ 是全连接层后的输出。

[0104] 网络训练算法如下:

[0105] <1>首先获取肺部CT图像公开数据集,选择其中的训练集作为训练模型的样本数据。

[0106] <2>将所有样本进行肺实质分割,获取要输入到模型中的样本。

[0107] <3>图像分割,得到的图像patches序列化输入到模型,并在初始化位置嵌入一个可学习的向量PE。

[0108] <4>图像序列化向量以及位置向量的结合输入到模型中,通过特征提取以及解码过程,输出肺结节的候选结果,即疑似肺结节区域的坐标、肺结节半径以及是否为肺结节的概率。

[0109] #循环,1~若干次

[0110] 根据<3>中的样本序列向量和位置向量PE,输入到网络中,使用随机梯度下降法计算判别网络参数:

$$[0111] \quad \text{Loss} = 1 - \frac{2|X \cap Y|}{|X| + |Y|}$$

[0112]  $X$ 是预测结果,对应预测出的mask, $Y$ 是Ground truth

[0113] 使用随机梯度下降法计算生成网络参数:

[0114] #循环结束

[0115] <5>算法结束。

[0116] 第四步、构建多尺度特征图。

[0117] 由于三维图像的计算量较为庞大,因此在构建特征图时,尽可能采取简单的策略。针对不同尺寸的肺结节,建立多种尺度的特征图。

[0118] 生成过程的具体算法如下:

[0119]  $F_1$ 生成 $L_1$ : $3 \times 3$ 卷积,尺寸为 $F_1$ 尺寸的1/2,即: $L_1 \in R^{N \times D/2 \times C_1}$ ,通道维度数为 $C_1$

$$[0120] \quad L_1 = \text{conv1}(F_1, C_1) \quad (15)$$

[0121]  $F_1$ 生成 $L_2$ : $3 \times 3$ 卷积,与 $F_1$ 尺寸相同,即: $L_2 \in R^{N \times D \times C_2}$ ,通道维度数为 $C_2$

$$[0122] \quad L_2 = \text{conv2}(F_1, C_2) \quad (16)$$

[0123]  $F_1$ 生成 $L_3$ : $3 \times 3$ 反卷积,尺寸为 $F_1$ 尺寸的两倍,即: $L_3 \in R^{N \times 2D \times C_3}$ ,通道维度数为 $C_3$

$$[0124] \quad L_3 = \text{deconv3}(F_1, C_3) \quad (17)$$



[0125]  $F_1$ 生成 $L_4:3 \times 3$ 反卷积,尺寸为 $F_1$ 尺寸的四倍,即: $L_4 \in R^{2N \times 2D \times C_4}$ ,通道维度数 $C_4$

[0126]  $L_4 = \text{deconv4}(F_1, C_4)$  (18)

[0127] 这四个尺度的特征输出是平行的,且彼此之间不存在自上而下或者自下而上的相连的路径。本发明特征图彼此之间没有连接或者融合计算,所以速度较快。在多尺度特征图上进行肺结节的检测,能够更多地检测出各种尺寸的肺结节,提高检测精度。

[0128] 第五步、肺实质图像稀疏化。

[0129] 由于三维图像包含内容较多,模型在计算时会受到很多背景信息或者无关信息的干扰,而且在肺部CT图像中,肺部的血管,孔洞容易给结节的检测带来干扰。本发明在上一步得到的P(肺实质图像分块后得到的patches)同时进行水平集算法处理得到包含肺结节或者其他孔洞组织的稀疏矩阵。肺结节或者其他内部组织,通过二值化图像确定它们在每一张肺实质切片图像内的轮廓如图4(b)、(c)所示,也就是一条二维平面内的闭合曲线。多个切片组合之后的曲线轮廓就构成了肺结节或者其他内部组织的表面曲线。这样可以在整个三维切片内部将所有疑似肺结节区域划分出来,原先的P就变成了一个稀疏矩阵。

[0130] 具体算法如下:

[0131] <1>读入任意一张肺部CT图像原图切片,图像灰度化,归一化,将它转化为可处理的数组类型;

[0132] <2>初始化水平集方程 $\text{IniLSF}$ ;

[0133] <3>根据水平集方程,画出初始轮廓,即利用 $\text{contour}$ 函数绘制等高线 $\text{LSF}=0$ 的轮廓线;

[0134] <4>定义CV(Chan-Vese)函数: $\text{CV}(\text{LSF}, \text{img}, \mu, \nu, \text{epison}, \text{step})$

[0135] 其中的参数有惩罚项系数 $\mu$ ,长度约束系数 $\nu$ ,规则化参数 $\text{epison}$ ,步长 $\text{step}$ , $\text{LSF}$ 是水平集等高线, $\text{img}$ 是待处理的图像

[0136] <5>#循环,迭代次数: $\text{iterNum}$

[0137] ①根据初始化的水平集计算(估算)前景和背景灰度均值;

[0138] ②利用前面估算的前景和背景均值对水平集每一个点进行调整(演化),根据定义前景 $>0$ ,背景 $<0$ ,如果当前点的灰度值接近前景估值,那么该点水平集的值相应的就要增加,反之则减小;

[0139] ③更新水平集方程

[0140] #循环结束;

[0141] <7>算法结束。

[0142] 本发明通过肺实质图像排除肺实质图像多余的干扰,方便接下来三视图自注意力机制的运算,同时简化卷积并节省乘积计算时的计算量。

[0143] 第六步、三视图自注意力机制计算

[0144] 将P的三种视图(正视图、侧视图、俯视图)分别输入网络,从多个视角进行联合学习。如图5所示,C(coronal)、A(axial)、S(sagittal)分别代表P的三个视角:正视图、俯视图、侧视图,它们的维度都为 $p \times p \times c$ 。运用 $3 \times 3$ 的卷积操作,将三个视角的特征投影到 $Q_{\text{query}}$ , $K_{\text{key}}$ 以及 $V_{\text{value}}$ 的维度:

[0145]  $Q_{\text{query}} = \text{conv}(A)$

[0146] Key=conv(S)

[0147] Value=conv(C) (19)

[0148] 其中,  $Q(\text{Query}) \in R^{H_q \times W_q \times D}$ ,  $K(\text{Key}) \in R^{H_k \times W_k \times D}$ ,  $V(\text{Value}) \in R^{H_v \times W_v \times D}$ 。

[0149] 下一步,将Q,K,V分别展平到二维空间,即:

$$Q = \text{Flatten}(Q), \quad Q: R^{H_q \times W_q \times D} \rightarrow R^{n \times D}$$

[0150] 
$$K = \text{Flatten}(K), \quad K: R^{H_k \times W_k \times D} \rightarrow R^{m \times D} \quad (20)$$

$$V = \text{Flatten}(V), \quad V: R^{H_v \times W_v \times D} \rightarrow R^{m \times D}$$

[0151] 根据self-attention公式,计算注意力权重W:

[0152] 
$$W = \text{softmax}\left(\frac{Q \times K^T}{\sqrt{D}}\right) \quad (21)$$

[0153] 其中,  $W \in [0, 1]^{n \times m}$ 。

[0154] 最终输出的特征为自注意力的权重W与V的乘积,即:

[0155] 
$$F_{sa} = W \times V \quad (22)$$

[0156] 然后将输出的 $F_{sa}$ 输入到前馈神经网络中来得到最后的特征输出 $F_2 \in R^{n \times D}$ 。

[0157] 第七步、使用训练好的网络模型进行肺结节的检测,输出预测结果

[0158] 图7为肺部CT图像的单张切片的去假阳前后的检测结果对比,a图为没有去假阳的检测结果,白色框内是准确的检测结果,灰色框内是误判的假阳性区域,b图为加入去假阳模块之后的结果,可以看出假阳性区域被筛除掉。

[0159] 本发明采用快速3D ViT网络结构提取特征图像信息,使用单尺度的输出依次构造多尺度特征图,解决了三维图像的数据量庞大的问题;3D肺实质图像包含大量背景点,只有小部分是有助于检测结果的重要信息。在结合多视角注意力的训练过程中,大量的背景点会给注意力模块带来意想不到的噪音。将分块后的图像稀疏化处理,即去掉大部分背景区域,可以凸显出肺结节所在的区域。使用自注意力机制提取肺实质图像的三视角信息用于假阳性区域的筛除。

[0160] 以上所述,仅是本发明的较佳实施例而已,并非是对本发明作其它形式的限制,任何熟悉本专业的技术人员可能利用上述揭示的技术内容加以变更或改型为等同变化的等效实施例应用于其它领域,但是凡是未脱离本发明技术方案内容,依据本发明的技术实质对以上实施例所作的任何简单修改、等同变化与改型,仍属于本发明技术方案的保护范围。

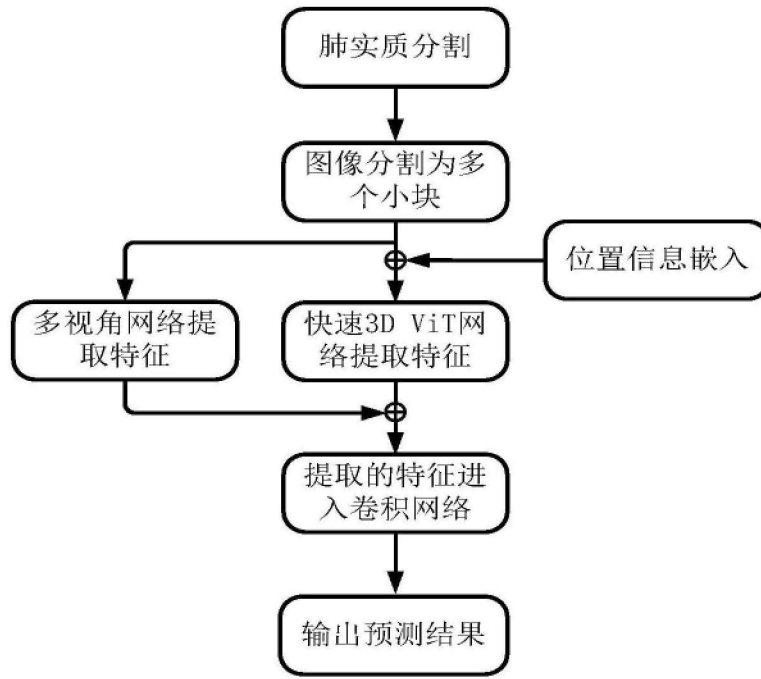


图1

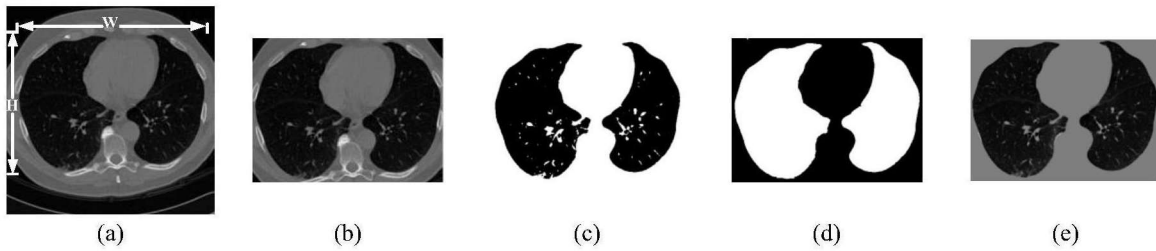


图2

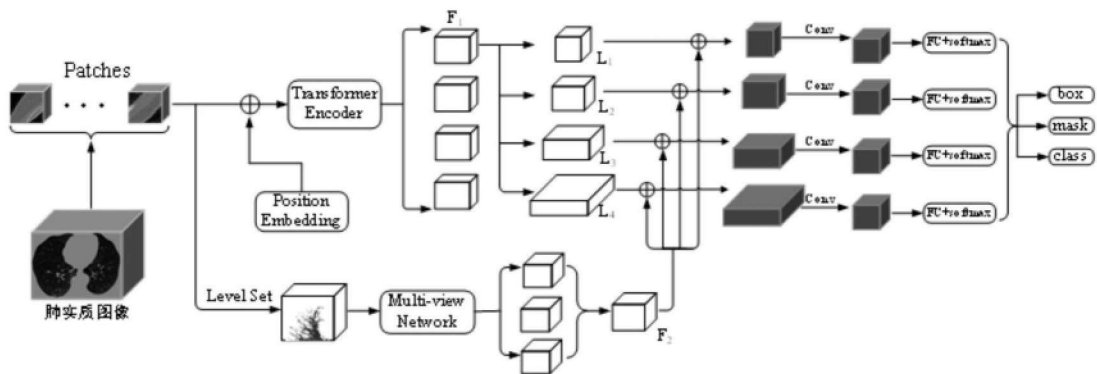


图3

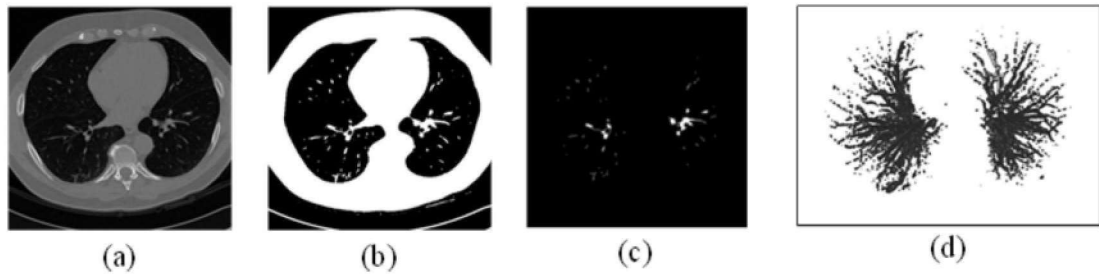


图4

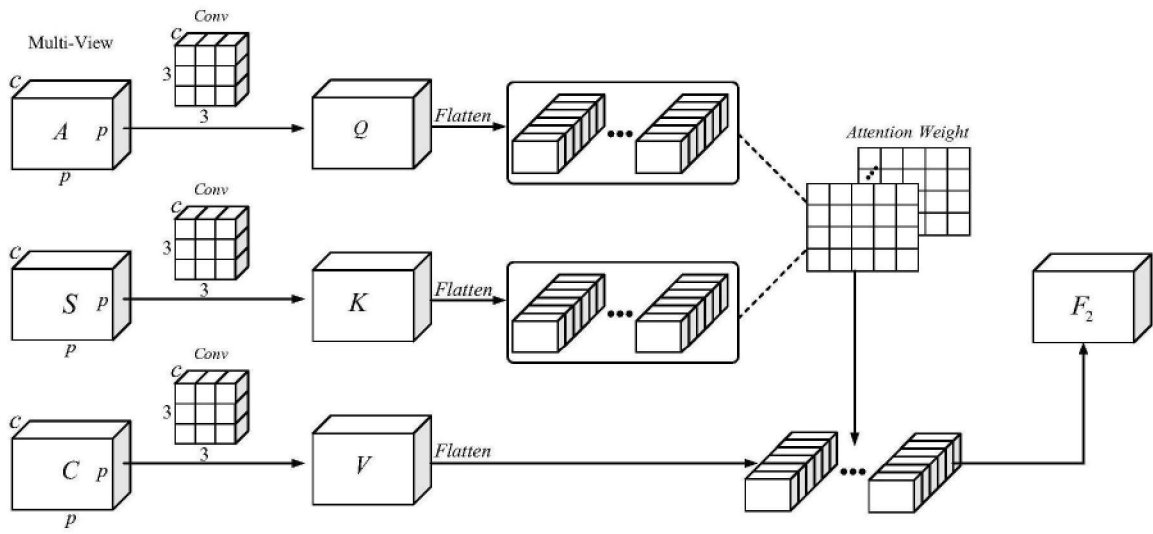


图5



图6

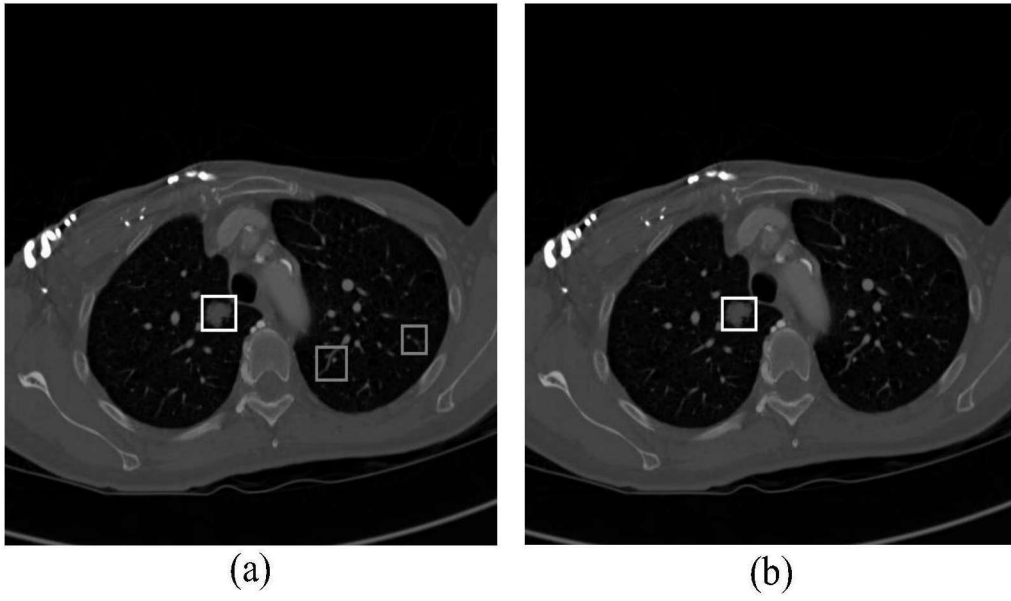


图7