



(12) 发明专利申请

(10) 申请公布号 CN 112529042 A

(43) 申请公布日 2021.03.19

(21) 申请号 202011294799.1

(22) 申请日 2020.11.18

(71) 申请人 南京航空航天大学

地址 210016 江苏省南京市秦淮区御道街
29号

(72) 发明人 张道强 朱文勇 孙亮

(74) 专利代理机构 南京经纬专利商标代理有限公司 32200

代理人 陆烨

(51) Int. Cl.

G06K 9/62 (2006.01)

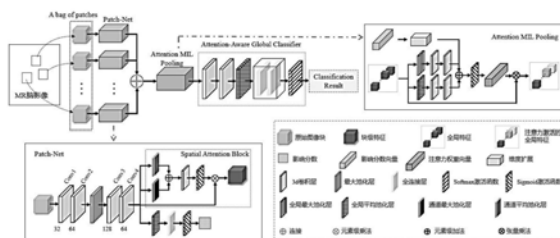
权利要求书3页 说明书9页 附图2页

(54) 发明名称

一种基于双重注意力多示例深度学习的医学图像分类方法

(57) 摘要

本发明公开了一种基于双重注意力多示例深度学习的医学图像分类方法,该方法包含如下步骤:1)采集M个医学图像,并基于分类任务对原始图像设置正负标签;2)基于医学图像的体素特征,确定K个固定位置作为输入图像特征的采样点;3)对双重注意力的多示例深度学习网络模型进行训练;4)更新多示例学习模型;5)将新的医学图像输入至训练好的多示例学习模型中,判断该医学图像的类别标签。本发明能够增强局部辨识度并权衡不同区域的特征信息权重,更符合医学图像的实际情况图像,更有助于提高医学图像分类性能。



1. 一种基于双重注意力多示例深度学习的医学图像分类方法,其特征在于,该方法包括如下步骤:

步骤1:采集M个医学图像,判断每个医学图像是否存在与特定任务相关的特征,将存在相关特征的医学图像的原始标签设置为正,将不存在相关特征的医学图像的原始标签设置为负;

步骤2:将原始标签为正的医学图像以及原始标签为负的医学图像都均匀划分为r个立方体;该r个立方体尺寸均为 $W*W*W$,将一个立方体作为一个图像块,基于原始标签为正的医学图像中所有图像块的体素特征和原始标签为负的医学图像中所有图像块的体素特征,在原始标签为正的医学图像中选择K个图像块,将该K个图像块所在的位置作为K个固定位置,提取M个医学图像中每个图像在K个固定位置上的图像块;

步骤3:将M个医学图像中的K个图像块作为多示例学习模型的输入,对多示例学习模型进行训练;

步骤4:计算每个医学图像的训练结果和该图像原始标签之间的交叉熵,建立基于交叉熵的损失函数,根据该损失函数判断训练后的多示例学习模型是否收敛,若是,则停止计算,得到训练好的多示例学习模型,否则根据损失函数计算得到的损失更新多示例学习模型的权重,并转步骤2;

步骤5:将新的医学图像输入至训练好的多示例学习模型中,判断该医学图像的分类。

2. 根据权利要求1所述的一种基于双重注意力多示例深度学习的医学图像分类方法,其特征在于,所述步骤2具体为:选择一组原始标签为正的医学图像作为正类图像组,选择一组原始标签为负的医学图像作为负类图像组,且两组图像组中医学图像的个数相同;根据正类图像组中每个图像的第b个图像块对应的体素信息,计算得到正类图像组中第b个图像块的平均值,根据负类图像组中每个图像的第b个图像块对应的体素信息,计算得到负类图像组中第b个图像块的平均值,将该两个平均值作为一组图像块组,从而组成对应于第b个图像块所在位置的两个特征平均值向量,采用t检验比较该两个特征平均值向量的差异性,得到第b个t检验结果,其中 $b=1,2,\dots,r$;对每一个t检验结果进行归一化计算,并将计算结果按照由小到大排列,选择前K个图像块组。

3. 根据权利要求1所述的一种基于双重注意力多示例深度学习的医学图像分类方法,其特征在于,所述步骤3中多示例学习模型的训练过程为:

步骤3.1:从输入的第m个医学图像的第k个图像块中提取影响分数和局部块级别的特征,其中 $m=1,2,\dots,M,k=1,2,\dots,K$;

步骤3.2:根据提取的K个局部块级别的特征和K个影响分数,计算得到m个医学图像的全局特征;

步骤3.3:根据全局特征计算第m个医学图像的分类。

4. 根据权利要求3所述的一种基于双重注意力多示例深度学习的医学图像分类方法,其特征在于,所述步骤3.1具体为:

步骤3.1.1:采用图像块级子网从输入的第m个医学图像的第k个图像块中提取特征 F , $F=\{F_1,F_2,\dots,F_c\}$, F_c 表示第c个特征,c为特征的总个数;所述图像块级子网包括依次连接的第一3D卷积层,第一批量归一化模块,第一激活函数,第二3D卷积层,第二批量归一化模块,第二激活函数,最大池化层,第三3D卷积层,第三批量归一化模块,第三激活函数,第四3D卷

积层,第四批量归一化模块以及第四激活函数;所述第一3D卷积层的卷积核为4*4*4,第二~四3D卷积层的卷积核均为3*3*3,所述最大池化层的大小为2*2*2;

步骤3.1.2:采用空间注意力模块提取F的局部块级别的特征;采用全局平均池化模块提取F的影响分数,所述全局平均池化模块包括依次连接的全连接层和第一sigmoid函数层。

5.根据权利要求4所述的一种基于双重注意力多示例深度学习的医学图像分类方法,其特征在于,所述步骤3.1.2中空间注意力提取F的局部块级别的特征具体为:

步骤A:采用通道最大池化层提取F的最大特征值 F_{max} :将F映射至第k个图像块的所有坐标位置中,得到F中的每个特征在坐标 (ω, h, l) 上的映射特征值并进行比较,从而得到该 (ω, h, l) 坐标上最大映射特征值 $F_{max}^{\omega, h, l}$,其中 $\omega = 0, 1, \dots, W, h = 0, 1, \dots, W, l = 0, 1, \dots, W$,将所有的 $F_{max}^{\omega, h, l}$ 组合成最大特征值 F_{max} ;

采用通道平均池化层提取F的平均特征值 $F_{average}$:将F映射至第k个图像块的所有坐标位置中,得到F中的每个特征在坐标 (ω, h, l) 上的映射特征值,并计算该 (ω, h, l) 坐标上所有特征值的平均值 $F_{average}^{\omega, h, l}$,将所有的 $F_{average}^{\omega, h, l}$ 组合成平均特征值 $F_{average}$;

步骤B:采用第五3D卷积层连接 F_{max} 和 $F_{average}$,并采用第二sigmoid函数层将连接结果限制在0到1范围内,从而得到空间注意力映射 $A_{spatial}$;

步骤C:将F中每个特征与 $A_{spatial}$ 逐元素相乘,得到局部块级别的特征 F_{attn} 。

6.根据权利要求3所述的一种基于双重注意力多示例深度学习的医学图像分类方法,其特征在于,所述步骤3.2具体为:将K个局部块级别的特征分别输入至K个通道方向的平均池化层,将该K个通道方向的平均池化层的输出结果进行组合得到全局特征 F_{global} ;

采用全局平均池化层提取 F_{global} 中全局平均特征向量,并依次采用第六3D卷积层和第七3D卷积层对全局平均特征向量进行计算,得到平均注意力权重映射 $A_{average}$:

$$A_{average} = w_1 \text{ReLU}(w_0 \text{GAP}(F_{global}))$$

w_0, w_1 分别是第六3D卷积层和第七3D卷积层的参数;ReLU(.)为ReLU激活函数,GAP(.)是全局平均池化函数,其中第六3D卷积层和第七3D卷积层的卷积核均为 $1 * 1 * 1$;

采用全局最大池化层提取 F_{global} 中全局最大特征向量,并依次采用第六3D卷积层和第七3D卷积层对全局最大特征向量进行计算,得到全局最大注意力权重映射 A_{max} :

$$A_{max} = w_1 \text{ReLU}(w_0 \text{GMP}(F_{global}))$$

其中,GMP(.)是全局最大池化函数;

将K个影响分数组合成一个向量a,将 $A_{average}, A_{max}$ 以及向量a进行组合,得到综合性注意力权重映射,采用第三sigmoid函数层将综合性注意力权重映射限制在0到1范围内,再将 F_{global} 中的元素与第三sigmoid函数层的结果进行逐相相乘,得到不同加权的全局特征。

7.根据权利要求3所述的一种基于双重注意力多示例深度学习的医学图像分类方法,其特征在于,所述步骤3.3中采用注意力激活的全局分类器计算第m个医学图像的正负类别的概率,所述注意力激活的全局分类器包括依次连接的第八3D卷积层,第九3D卷积层,全局平均池化层,第一全连接层,第二全连接层和softmax激活函数。

8.根据权利要求1所述的一种基于双重注意力多示例深度学习的医学图像分类方法,其特征在于,所述步骤4中,采用Adam优化函数根据损失函数计算得到的损失更新多示例学

习模型的权重。

一种基于双重注意力多示例深度学习的医学图像分类方法

技术领域

[0001] 本发明属于图像分类技术领域。

背景技术

[0002] 医学图像具有高维和高尺寸的特点,但是医学图像含有大多数的无关背景信息,对最后的分类判断造成严重的干扰。例如在三维的结构核磁共振(MR)图像判断局部微小的脑萎缩现象,在眼底视网膜图像上判断血管出血现象以及在组织图像中判断是否含有某个特定的细胞等。这些医学图像在不同类别上显示全局相似性,只有局部的差异,直接在完整图像上进行计算机分析往往性能较差。

[0003] 传统的基于机器学习的医学图像分类方法通常将医学图像划分成多个区域进行分析,来克服在缺少医学先验知识的前提下直接分析完整图像的困难。依据划分的区域特征表示大小,目前的医学图像分类研究可以分为三类:1)体素级别(voxel-level),2)区域级别(region-level),和3)图像块级(patch-level)方法。对于体素级方法,一般通过从医学影像提取所有体素级组织特征,来组成非常高维特征向量作为分类模型的输入特征。然而与庞大的特征数目相比,用于图像分类的训练图像往往非常少,由于医学影像的珍贵性和被试个体的稀有性。所以这类方法往往面临维度灾难(Curse of Dimensionality)和过拟合等问题。为了降低特征维度,区域级方法则基于从医学图像中分割的感兴趣区域(ROI)来进行分类预测,通过从ROI中提取某种生物特征来区分医学图像的类别。而这类方法是需要专家经验知识和专门的工具来手工提取特征的,是低成本和费时的。并且特征提取过程独立于之后的分类算法,两者之间存在异质性,容易造成次优的分类性能。而最近提出的块级方法被证明有效地描述医学图像中的局部结构变化。它是一种体素级和区域级之间的中间大小规模,更为灵活。并且图像块的中心位置可以通过特定数据驱动方法和统计学方法给出建议。然而,块级方法也面临许多挑战。比如块级标签定义模糊问题,不能直接用图像级标签定义图像块的标签。多示例学习作为一种弱监督学习被用来解决这个问题,但同时也存在很多问题。如Tong基于图像块特征值的欧几里得距离,以距离矩阵,构建图核作为全局特征,最后利用支持向量机进行分类。这种方法未考虑图像块内的空间信息。Liu等构建多个子卷积神经网络(sub-CNN)学习块级特征表示,并将学得的块级特征按通道简单连接,最后用全连接网络分类。这种无差别连接块级特征组成全局特征是不合理的。因为医学图像中不同图像块含有的信息量权重是不同的。块级方法仍需要解决的主要难题在于如何组合局部的块级特征为一个可以代表整个图像信息的全局特征表示。

[0004] 以上传统医学图像分类方法存在三个明显的缺陷:1)忽视了各个被划分的图像区域内的空间信息,和缺少对具有辨识力的结构部分的特征增强,这样容易造成不同个体特征差异过低从而分类性能较差。2)只提取目标区域的部分生物特征,可能未包含可用于图像分类的全部相关特征;只利用部分独立的区域特征,不能代表个体的整体特征。3)无差别地组合局部区域的特征表示,然而不同的图像区域含有的信息量是不同,有些包含无关的信息甚至成为分类噪声。所以对信息量高的图像区域的特征增强和对无关的图像区域的

特征减弱是非常重要的。

[0005] 近来深度学习方法在医学图像分类方面表现较高的性能,相比与传统学习使用手工提取的特征,深度学习模型可以以纯数据驱动的方式自动学得任务相关的特征,从而消除特征学习与分类算法的异质性,提高了分类的性能。但是由于深度方法中神经网络的黑盒性质,目前的深度学习方法都缺少与分类任务相关的特征的定位的反馈。

发明内容

[0006] 发明目的:为解决上述背景技术中存在的问题,本方发明提供了一种用于医学图像分类的双重注意力多示例深度学习模型。

[0007] 技术方案:一种用于医学图像分类的双重注意力多示例深度学习模型,其特征在于:该模型包括如下步骤:

[0008] 步骤1:采集M个医学图像,并根据每个医学图像是否存在特定任务相关特征,(如MR脑图像某个脑区是否发生异常脑萎缩,组织图像是否含有某个特定细胞),对每个医学图像设置原始标签,分为正负两类;

[0009] 步骤2:将正类的医学图像以及负类的医学图像都均匀划分为r个立方体;该r个立方体尺寸均为 $W*W*W$,将一个立方体作为一个图像块,基于正类的图像中所有图像块的体素特征和负类的图像中所有图像块的体素特征,在医学图像中选择K个图像块,将该K个图像块所在的位置作为K个固定位置,提取m个医学图像中每个图像在K个固定位置上的图像块;

[0010] 步骤3:将M个医学图像中的K个图像块作为多示例学习模型的输入,对多示例学习模型进行训练;

[0011] 步骤4:计算每个医学图像的训练结果和该图像原始标签之间的交叉熵,建立基于交叉熵的损失函数,根据该损失函数判断训练后的多示例学习模型是否收敛,若是,则停止计算,得到训练好的多示例学习模型,否则根据损失函数计算得到的损失更新多示例学习模型的权重,并转步骤2;

[0012] 步骤5:将新的医学图像输入至训练好的多示例学习模型中,判断该医学图像的类别。

[0013] 进一步的,所述步骤2具体为:选择一组正类的医学图像,和一组负类的医学图像,两组中医学图像的个数相同;根据该组正类图像中每个图像的第b个图像块对应的体素信息,计算得到正类第b个图像块的平均值,根据该组负类图像中每个图像的第b个图像块对应的体素信息,计算得到负类的第b个图像块的平均值,将该两个平均值作为一组图像块组,从而组成对应于第b个图像块所在位置的两组特征平均值向量,采用t检验比较该两组特征平均值向量的差异性,得到第b个t检验结果,其中 $b=1,2,\dots,r$;对每一个t检验结果进行归一化计算,并将计算结果按照由小到大排列,选择前K个图像块组。

[0014] 进一步的,所述步骤3中多示例学习模型的训练过程为:

[0015] 步骤3.1:从输入的第m个医学图像的第k个图像块中提取影响分数和局部块级别的特征,其中 $m=1,2,\dots,M,k=1,2,\dots,K$;

[0016] 步骤3.2:根据提取的K个局部块级别的特征和K个影响分数,计算得到m个医学图像的全局特征;

[0017] 步骤3.3:根据全局特征计算第m个医学图像的类别。

[0018] 进一步的,所述步骤3.1具体为:

[0019] 步骤3.1.1:采用块级子网从输入的第m个医学图像的第k个图像块中提取特征 F , $F = \{F_1, F_2, \dots, F_c\}$, F_c 表示第c个特征, c为特征的总个数;所述块级子网包括依次连接的第一3D卷积层,第一批量归一化模块,第一激活函数,第二3D卷积层,第二批量归一化模块,第二激活函数,最大池化层,第三3D卷积层,第三批量归一化模块,第三激活函数,第四3D卷积层,第四批量归一化模块以及第四激活函数;所述第一3D卷积层的卷积核为 $4 \times 4 \times 4$,第二~四3D卷积层的卷积核均为 $3 \times 3 \times 3$,所述最大池化层的大小为 $2 \times 2 \times 2$;

[0020] 步骤3.1.2:采用全局平均池化模块提取F的影响分数,采用空间注意力模块提取F的局部块级别的特征;所述全局平均池化模块包括依次连接的全连接层和第一 sigmoid函数层。

[0021] 进一步的,所述步骤3.1.2中空间注意力提取F的局部块级别的特征具体为:

[0022] 步骤A:采用通道最大池化层提取F的最大特征值 F_{\max} 具体为:将F映射至第k个图像块的所有坐标位置中,得到F中的每个特征在坐标 (ω, h, l) 上的映射特征值并进行比较,从而得到该 (ω, h, l) 坐标上最大映射特征值 $F_{\max}^{\omega, h, l}$,其中 $\omega = 0, 1, \dots, W, h = 0, 1, \dots, W, l = 0, 1, \dots, W$,将所有的 $F_{\max}^{\omega, h, l}$ 组合成最大特征值 F_{\max} ;

[0023] 采用通道平均池化层提取F的平均特征值 F_{average} 具体为:将F映射至第k个图像块的所有坐标位置中,得到F中的每个特征在坐标 (ω, h, l) 上的映射特征值,并计算该 (ω, h, l) 坐标上所有特征值的平均值 $F_{\text{average}}^{\omega, h, l}$,将所有的 $F_{\text{average}}^{\omega, h, l}$ 组合成平均特征值 F_{average} ;

[0024] 步骤B:采用第五卷积层连接 F_{\max} 和 F_{average} ,并采用第二sigmoid函数层将连接结果限制在0到1范围内,从而得到空间注意力映射 A_{spatial} ;

[0025] 步骤C:将F中每个特征与 A_{spatial} 逐元素相乘,得到局部块级别的特征 F_{attn} 。

[0026] 进一步的,所述步骤3.2具体为:将K个局部块级别的特征分别输入至K个通道方向的平均池化层,将该K个通道方向的平均池化层的输出结果进行组合得到全局特征 F_{global} ;

[0027] 采用全局平均池化层提取 F_{global} 中全局平均特征向量,并依次采用第六卷积层和第七卷积层对全局平均特征向量进行计算,得到平均注意力权重映射 A_{average} :

[0028] $A_{\text{average}} = w_1 \text{ReLU}(w_0 \text{GAP}(F_{\text{global}}))$

[0029] w_0, w_1 分别是第六卷积层和第七卷积层的参数; $\text{ReLU}(\cdot)$ 为ReLU激活函数, $\text{GAP}(\cdot)$ 是全局平均池化函数,其中第六卷积层和第七卷积层的卷积核均为 $1 \times 1 \times 1$;

[0030] 采用全局最大池化层提取 F_{global} 中全局最大特征向量,并依次采用第六卷积层和第七卷积层对全局最大特征向量进行计算,得到全局最大注意力权重映射 A_{max} :

[0031] $A_{\text{max}} = w_1 \text{ReLU}(w_0 \text{GMP}(F_{\text{global}}))$

[0032] 其中, $\text{GMP}(\cdot)$ 是全局最大池化函数;

[0033] 将K个影响分数组合成一个向量a,将 $A_{\text{average}}, A_{\text{max}}$ 以及向量a进行组合,得到综合性注意力权重映射,采用第三sigmoid函数层将综合性注意力权重映射限制在0到1范围内,再将 F_{global} 中的元素与第三sigmoid函数层的结果进行逐相相乘,得到不同加权的全局特征。

[0034] 进一步的,所述步骤3.3中采用注意力激活的全局分类器计算第m个医学图像的正负类别的概率,所述注意力激活的全局分类器包括依次连接的第八3D卷积层,第九3D卷积

层,全局平均池化层,第一全连接层,第二全连接层和softmax激活函数。

[0035] 进一步的,所述步骤4中,采用Adam优化函数根据损失函数计算得到的损失更新多示例学习模型的权重。

[0036] 有益效果:

[0037] 1、本发明可自动学得医学图像分类的相关特征,不需要手工的感兴趣区域特征提取过程,很大程度上提高了时间效率,实现医学图像分类的实时性,有助于辅助医生临床诊断。本发明中的多示例习模型由多个带有空间注意力模块的块级网络,一个注意力多示例池化模块和一个全局注意力激活的分类器构成。能够提取局部特征,并融合成代表整脑结构的全局特征,这种结构能够增强局部辨识力高的特征,并权衡不同区域的特征信息权重,更符合医学图像的实际情况,即在高维的医学图像中寻找微弱的病理特征变化,更有助于提高分类性能,增加了模型的可解释性。

[0038] 2、本发明中的块级子网络能够从医学图像中的多个图像块中提取局部特征,并且图像块的大小较像素级和区域级更加灵活,更适合学得局部特征。并且每个块级子网络可以获得每个块对最终分类判断的影响分数,有助于判别任务相关特征位置。

[0039] 3、本发明中的空间注意力模块为每个块级网络连接一个空间注意力模块,可以有效识别每个块内辨识度高的特征,即块内一些与分类任务高度相关的微小结构变化,并对这些特征进行增强。

[0040] 4、本发明中注意力多示例池化模块用于将块级子网络输出的局部特征进行加权融合,合并成一个全局的特征。与一般方法直接按通道连接不同,该模块可以为不同的图像块级局部特征赋予不同的权重,提高了输出结果的准确度,因为每个图像块含有的辨识力信息量是不同,有些图像块可能没有分类任务相关的特征,或者含有干扰因素,所以需要增强有关图像的特征权重,而削弱无关图像块的特征权重。

[0041] 5、本发明对于不同维度的(3D/2D)图像,采用相应维度(3D/2D)的卷积层结构,适用于在较大尺寸的图像中检测较小的目标,采用注意力机制和多示例学习可以减弱背景信息的干扰,找到关键的特征信息。

附图说明

[0042] 图1本发明的结构图;

[0043] 图2为不同的多示例池化在分类性能上的比较;

[0044] 图3不同图像块大小在分类性能上的结果;

[0045] 图4不同图像块数量在分类性能上的结果。

具体实施方式

[0046] 构成本发明的一部分的附图用来提供对本发明的进一步理解,本发明的示意性实施例及其说明用于解释本发明,并不构成对本发明的不当限定。

[0047] 如图1所示实施例提供了一种基于双重注意力多示例深度学习的医学图像分类方法,由于特征变化发生在局部区域,并非所有来自正类的sMRI图像的分割图像块都在形态上发生异常变化。这个现象正好符合多示例学习问题。在多示例学习中,训练数据是由一个包的集合组成,每个包包含有多个示例,只有包级别的标签是已知的,而示例的标签是未知

的。即, $D = \{(X_i, Y_i)\}_{i=1}^N$, 其中 $X_i = \{I_{ij}\}_{j=1}^{N_i}$, Y_i 是包 X_i 的标签, N 是包的个数, N_i 是包 X_i 中示例的个数。将一个正类的MR图像提取的图像块组成一个正标记的包, 在这个包中至少有一个正类的块, 即至少有一个图像块是受到影响的。将一个负类的图像中提取的块组成一个负标记包, 在这个负包中, 所有的图像块都是负类。而多示例学习模型DA-MIDL模型的任务就是利用多个图像块来实现对未知包(即未知图像)的正确分类。

[0048] 本实施例的方法包括如下步骤:

[0049] 步骤1: 采集M个医学图像, 判断每个医学图像是否存在与特定任务相关的特征, 将存在相关特征的医学图像的原始标签设置为正, 将不存在相关特征的医学图像的原始标签设置为负; 所述特定任务包括: MR图像中是否存在脑萎缩, 脑萎的种类, 或者组织图像中是否存在特定的细胞等;

[0050] 步骤2: 将原始标签为正的医学图像以及原始标签为负的医学图像都均匀划分为r个立方体; 该r个立方体尺寸均为W*W*W, 将一个立方体作为一个图像块, 基于原始标签为正的医学图像中所有图像块的体素特征和原始标签为负的医学图像中所有图像块的体素特征, 在原始标签为正的医学图像中选择K个图像块, 将该K个图像所在的位置作为K个固定位置, 提取M个医学图像中每个图像在K个固定位置上的图像块;

[0051] 步骤3: 将M个医学图像中的K个图像块作为多示例学习模型的输入, 对多示例学习模型进行训练;

[0052] 步骤4: 计算每个医学图像的训练结果和该图像原始标签之间的交叉熵, 建立基于交叉熵的损失函数, 根据该损失函数判断训练后的多示例学习模型是否收敛, 若是, 则停止计算, 得到训练好的多示例学习模型, 否则根据损失函数计算得到的损失更新多示例学习模型的权重, 并转步骤2;

[0053] 步骤5: 将新的医学图像输入至训练好的多示例学习模型中, 判断该医学图像的类别。

[0054] 在本实施例中, 由于需要很高的计算和存储成本把医学图像的所有图像块输入到网络中分析, 所以图像块位置的对于提取辨识度信息是必不可少。我们提出了一种新的图像块位置建议方法。首先按照不重叠的固定位置, 将医学图像均匀划分为具有固定大小的立方体, 以简化计算, 避免冗余信息。并不是所有的分割块都与分类任务有关, 这意味着每个图像块含有不同的信息量。因此, 我们应用一种统计方法对整个图像中所有块的信息量进行排序。我们计算了一个块中所有体素特征的平均值作为其块级别特征。然后, 使用t检验对两组分别来自训练集中正类和负类的块级特征进行了差异性比较。所以我们可以在这个图像块位置得到一个p-value, 它可以代表这个位置的信息性。此外, p-value越小, 图像块位置的辨识度越高。对所有位置上的p-value值进行归一化, 可以形成覆盖整个医学图像的p-value图, 并将归一化后的结果按照由小到大排列, 选择前K个图像块组, 组成多示例学习的一个包, 作为模型的训练样本

[0055] 在本实施例中, 所述多示例学习模型(DA-MIDL)的学习过程具为: 从M个医学图像中的一个医学图像中选择K个图像块, 采用带有空间注意力模块的图像块级子网(Patch-Net with Spatial Attention Block)从第k个图像块中提取影响分数Affect score和局部块级别的特征, 然后将将块级别的特征表示联合并输入到注意力多示例池化中, 计算每个块的相对重要性, 最后通过全局分类器对注意感知的全局特征进行进一步的学习和分

类。

[0056] 在本实施例中,带有空间注意力模块的块级子网包括块级子网Patch-Net、空间注意力模块和全局平均池化模块;所述Patch-Net从原始图像块中学习更加抽象的特征表示 F , $F = \{F_1, F_2, \dots, F_c\}$, F_c 表示第 c 个特征, c 为特征的总个数,并且减少feature map的大小;所述块级子网包括依次连接的第一3D卷积层,第一批量归一化模块,第一激活函数,第二3D卷积层,第二批量归一化模块,第二激活函数,最大池化层,第三3D卷积层,第三批量归一化模块,第三激活函数,第四3D卷积层,第四批量归一化模块以及第四激活函数;所述第一3D卷积层的卷积核为 $4 \times 4 \times 4$,第二~四3D卷积层的卷积核均为 $3 \times 3 \times 3$,所述最大池化层的大小为 $2 \times 2 \times 2$;2个单位长度的步长为了向下采样,第一~四卷积核通道数分别是32,64,128和64;基于第四卷积层conv4输出的feature maps,块级子网延伸出两个分支模块,一个是空间注意力模块:用于学习一个空间注意力的局部块级特征。另一个模块是一个全局平均池化,包括依次连接的全连接层和第一sigmoid函数组成。为了输出一个影响分数,可以帮助识别可能的病灶位置。与其他多示例学习的示例级别特征转化不同,从块级网络输出的块级特征表示任然保留三维的形状而不是一维特征向量,是为了之后块级特征表示更好的结合和进一步学习图像块之间的关系。

[0057] 在本实施例中,空间注意力模块学习一个空间注意力的局部块级特征具体为:为了增强局部辨识力结构特征,设计了一个空间注意力模块嵌入到块级网络中。两种沿着通道轴的不同池化(即通道最大池化和通道平均池化)被采用于生成两个特征映射分别代表最大特征和平均特征。然后这两个特征映射被联合输入到之后的第五3D卷积层。该层卷积的输出可以被看作是一个空间注意力映射 $A_{spatial}$,它的大小等于conv4输出的特征映射的大小。每个位置的注意力分数通过第二sigmoid函数限制在0到1的范围内。这个空间注意力地图可以描述图像块内各个部分空间变化的贡献度。然后conv4输出的每个特征映射与空间注意力映射逐元素相乘,最终生成一个局部的带空间注意力的特征表示 F_{attn} :

$$[0058] \quad F_{attn} = [F_1 \otimes A_{spatial}, \dots, F_c \otimes A_{spatial}]$$

[0059] 其中, \otimes 表示逐元素相乘。

[0060] 采用通道最大池化层提取 F 的最大特征值 F_{max} 具体为:将 F 映射至第 k 个图像块的所有坐标位置中,得到 F 中的每个特征在坐标 (ω, h, l) 上的映射特征值并进行比较,从而得到该 (ω, h, l) 坐标上最大映射特征值 $F_{max}^{\omega, h, l}$,其中 $\omega = 0, 1, \dots, W, h = 0, 1, \dots, W, l = 0, 1, \dots, W$,将所有的 $F_{max}^{\omega, h, l}$ 组合成最大特征值 F_{max} ;

[0061] 采用通道平均池化层提取 F 的平均特征值 $F_{average}$ 具体为:将 F 映射至第 k 个图像块的所有坐标位置中,得到 F 中的每个特征在坐标 (ω, h, l) 上的映射特征值,并计算该 (ω, h, l) 坐标上所有特征值的平均值 $F_{average}^{\omega, h, l}$,将所有的 $F_{average}^{\omega, h, l}$ 组合成平均特征值 $F_{average}$ 。

[0062] 在本实施例中,注意力多示例池化用于学习一个块相对注意映射来表示每个图像块的相对权重;块级特征表示在连接成一个全局结构特征表示之前,经过通道方向的平均池化,使得通道数等于图像块的个数(将 K 个局部块级别的特征分别输入至 K 个通道方向的平均池化层,将该 K 个通道方向的平均池化层的输出结果进行组合得到全局特征 F_{global}),每个通道上的特征映射代表一个图像块。全局平均池化和最大池化平行构建用于生成两个

不同的特征描述。然后这两个特征描述被均被第六和第七卷积层学习产生两个patch-attentionmap:

$$[0063] \quad A_{\text{average}} = w_1 \text{ReLU}(w_0 \text{GAP}(F_{\text{global}}))$$

$$[0064] \quad A_{\text{max}} = w_1 \text{ReLU}(w_0 \text{GMP}(F_{\text{global}}))$$

[0065] w_0, w_1 分别是第六卷积层和第七卷积层的参数; $\text{ReLU}(\cdot)$ 为ReLU激活函数, $\text{GAP}(\cdot)$ 是全局平均池化函数, 其中第六卷积层和第七卷积层的卷积核均为 $1*1*1$; 其中, $\text{GMP}(\cdot)$ 是全局最大池化函数;

[0066] 特别的是, 在我们实验中处理最大特征描述和平均特征描述的卷积层享有相同的参数。除了这两个通过学习块间关系得到的patch-attention maps, 只通过单个块内特征得到的影响分数affect score也被考虑去评估每个图像块的相对权重。从每个Patch-Net得到的affect score可以组成一个向量 $a(a_1, a_2, \dots, a_c)$, 其中 a_c 为第 c 个影响分数, 这个向量被拓展为patch-attention map相同的大小, 然后这三个不同的注意力映射组合成一个更加全面的patch-attention map即 A_{patch} , 之后被第三sigmoid函数激活。

$$[0067] \quad A_{\text{patch}} = \sigma(A_{\text{average}} + A_{\text{max}} + a)$$

[0068] 其中 $\sigma(\cdot)$ 为sigmoid函数

[0069] 最后, 之前的全局特征表示与该注意力映射通过张量乘法得到一个注意力激活的全局特征:

$$[0070] \quad \mathcal{F}_{\text{global}} = \mathbf{F}_{\text{global}} \otimes A_{\text{patch}}$$

[0071] 与传统的最大多示例池化和平均多示例池化不同, 本是实施例中注意力多示例池化 (Attention MIL Pooling) 不仅考虑全部图像块的特征而不是只关注部分辨识力高的图像块, 并且赋予每个图像块不同的权重而不是无差别的组合图像块。因此, 这个注意力多示例池化可以增强关键图像块的特征从而减少噪声干扰, 同时保留了不重要的图像和关键图像的潜在的关联信息从而降低了特殊个体的误判率。另外, 计算得到的 patch-attention map可以作为是特征定位的参考。

[0072] 在本实施例中, 注意力激活的全局分类器 (Attention-Aware Global Classifier) 通过分析块间的高度相关性继续学习全局特征表示, 并作出最后的分类结果。因为相对于全连接层, 卷积层展示一个较好的非线性高维特征提取能力。所以全局分类器的前面两层卷积网络用来进一步学习注意激活的全局特征表示来获取更多的结构信息和压缩特征映射的大小。同时一个3D全局平均池化用于下采样。然后特征表示被展平作为之后两层全连接层的输入, 产生两个分数, 分别代表负类和正类的概率。

[0073] 基于不同加权的特征映射, 注意感知全局分类器, 以继续学习全局特征表示可代表MRI图像的全脑结构信息, 并输出AD检测或MCI转换预测的分类结果。

[0074] 在本是实施例中, 由于只有图像级别的标签是给出的, 而块级的标签是模糊的, 图像级别的标签信息被认为是用于反向传播更新我们的网络权重 w 的唯一指导。采用基于交叉熵损失函数的损失函数描述为 $\mathfrak{L}(w)$:

$$[0075] \quad \mathfrak{L}(w) = -\frac{1}{M} \sum_{m=1}^M \log(P(Y_m | X_m; w))$$

[0076] 其中, $P(Y_m | X_m; w)$ 是当前网络在权重为 w 下正确预测的概率。

[0077] 在本实施例中,除了医学图像分类,本实施例提供的DA-MIDL模型还可以自动识别特征的位置。

[0078] 以MR图像为例进行仿真实验:

[0079] 实验数据集是从ADNI公开数据库获取的。总共包含了由1193个被试的baseline 1.5T T1-weighted的结构性MR图像。这些图像可以分为三个类别:重度脑萎缩,轻度脑萎缩和正常对照。轻度脑萎缩对象可以进一步被分为两类:进展性脑萎缩(即36个月内转化为重度脑萎缩)和稳定性脑萎缩(即一直稳定在轻度)。我们研究的数据集总共包含了389个重度脑萎缩,172个进展性脑萎缩,232个稳定性脑萎缩和400个正常对照。

[0080] 图像预处理:将每个MR图像与Colin27模板线性配准消除全局线性差异(全局平移、尺度和旋转差异),然后去除图像中的头颅部分。最后得到的图像大小为 $181 \times 217 \times 181$ 。

[0081] 实验设定:

[0082] 本实施例在多个分类任务上得到了验证。即重度脑萎缩和正常对照分类,进展性脑萎缩与稳定性脑萎缩分类,进展性脑萎缩和正常对照分类,稳定性脑萎缩和正常对照分类。并使用5折交叉验证法。

[0083] 使用四个指标验证分类性能,包括准确率(ACC),敏感性(SEN),特异性(SPE)和ROC曲线下的面积(AUC)。

[0084] MR脑影像分类结果:

[0085] 对采用本实施例的实验结果与四种方法进行了比较(体素级特征+SVM分类 Voxel-based Morphometry-VBM, ROI级特征+SVM分类 ROI-wise Method-ROI, 块级特征+深度多示例学习 Deep Multi-instance Learning-DMIL, 块级特征+全卷积网 Hierarchical Fully Convolutional Network-HFCN), 结果如表1和表2所示,其中表1为重度脑萎缩和正常对照分类,进展性脑萎缩与稳定性脑萎缩分类的实验结果,表2为进展性脑萎缩和正常对照,稳定性脑萎缩和正常对照分类的实验结果(Method为方法, ACC为准确率, SEN为敏感性, SPE为特异性, AUC为接收者操作特征曲线下的面积) 显然本实施例在四个分类任务上都取得了较好的结果。

[0086] 表1

Method	重度脑萎缩 vs. 正常对照				进展性脑萎缩 vs. 稳定性脑萎缩			
	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
VBM	0.822	0.815	0.829	0.858	0.673	0.614	0.716	0.703
ROI	0.803	0.748	0.857	0.837	0.669	0.589	0.729	0.699
DMIL	0.895	0.873	0.916	0.948	0.764	0.678	0.827	0.778
HFCN	0.901	0.889	0.912	0.954	0.781	0.650	0.878	0.816
DA-MIDL	0.930	0.938	0.926	0.961	0.803	0.843	0.773	0.830

[0088] 表2

Method	进展性脑萎缩 vs. 正常对照				稳定性脑萎缩 vs. 正常对照			
	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
VBM	0.819	0.682	0.879	0.836	0.638	0.596	0.663	0.675

[0089]

[0090]	ROI	0.773	0.737	0.789	0.824	0.616	0.577	0.638	0.627
	DMIL	0.878	0.865	0.883	0.929	0.798	0.765	0.817	0.808
	HFCN	0.874	0.869	0.877	0.926	0.802	0.738	0.839	0.817
	DA-MIDL	0.893	0.881	0.898	0.939	0.835	0.898	0.798	0.856

[0091] 从表1和表2中,可以得到以下推断。1) 与传统的方法(VBM和ROI)相比,本实施例在分类性能上有显著的提高。深度学习方法连接了特征提取和分类器学习过程,可以生成一个较好的面向特定分类任务的特征表示。2) 块级方法比其他体素级和区域级方法在结果上要好。表明块级的特征表示更能反应局部的脑萎缩形态特征。3) 与DMIL 和HFCN两种先进的方法比较,本实施例都取得较好的实验结果。在相同的输入特征信息下,本实施例的能学得一个更优越的特征表示来识别MR影像中脑萎缩位置。

[0092] 注意力模块的有效性验证:

[0093] 根据是否包含注意力模块,本实施例的DA-MIDL模型可以分为四类:1) 不包含任何注意力模块(N-MIL), 2) 只包含空间注意力模块(S-MIL), 3) 只包含注意力多示例池化(A-MIL)和4) 包含双重注意力模块(DA-MIDL)。在重度脑萎缩和正常对照分类,进展性脑萎缩与稳定性脑萎缩分类上进行实验验证,结果显示在表3中。

[0094] 表3

Method	重度脑萎缩 vs. 正常对照				进展性脑萎缩 vs. 稳定性脑萎缩			
	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
[0095] N-MIDL	0.879	0.875	0.883	0.927	0.740	0.681	0.783	0.765
S-MIDL	0.899	0.895	0.902	0.942	0.750	0.756	0.745	0.788
A-MIDL	0.908	0.928	0.890	0.956	0.774	0.802	0.754	0.795
DA-MIDL	0.930	0.938	0.926	0.961	0.803	0.843	0.773	0.830

[0096] 从结果上看,本实施例的空间注意力模块和注意力多示例池化,可以有效提升分类准确率。并且包含双重注意力模块的模型比只使用一种注意模块的模型性能要好。所以本实施例提出的注意力模块可以从MR图像中有效地捕捉辨识度高的特征。

[0097] 并且本实施例也将提出的注意力多示例池化(Attention MIL Pooling)与常见的多示例池化(即平均多示例池化Average MIL Pooling和最大多示例池化Max MIL Pooling)进行了比较。实验结果如图2,其纵坐标为分类性能。

[0098] 本实施例提出的注意力多示例池化可以有效提升分类性能。同时弥补了平均多示例池化和最大多示例池化的缺点,不仅生成一个全局的特征表示,而且对不同辨识力的特征赋予不同的权重。

[0099] 模型自身参数(图像块的大小和数量)的影响:

[0100] 考虑了输入图像块的大小和数量对模型性能的影响,取图像块的大小范围从 $15 \times 15 \times 15$ 到 $35 \times 35 \times 35$ 。图像块的数量范围{20, 40, 60, 80, 100, 120},同样也在分类性能上进行比较。实验结果如图3,图4,最后处于性能和计算成本的考虑,本实施例选择64个大小为 $25 \times 25 \times 25$ 的图像块作为模型的输入。

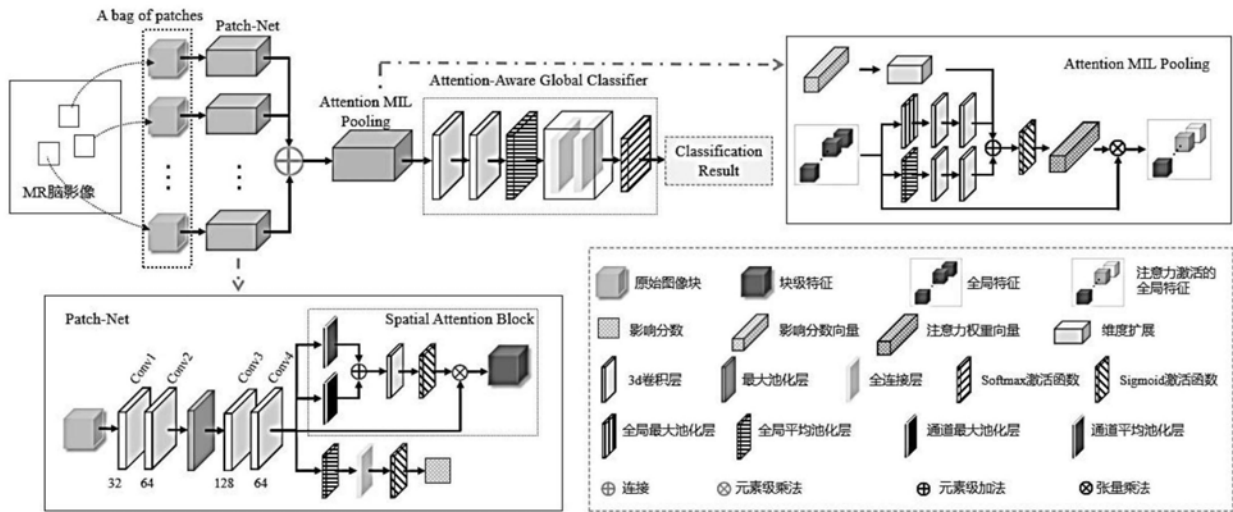


图1

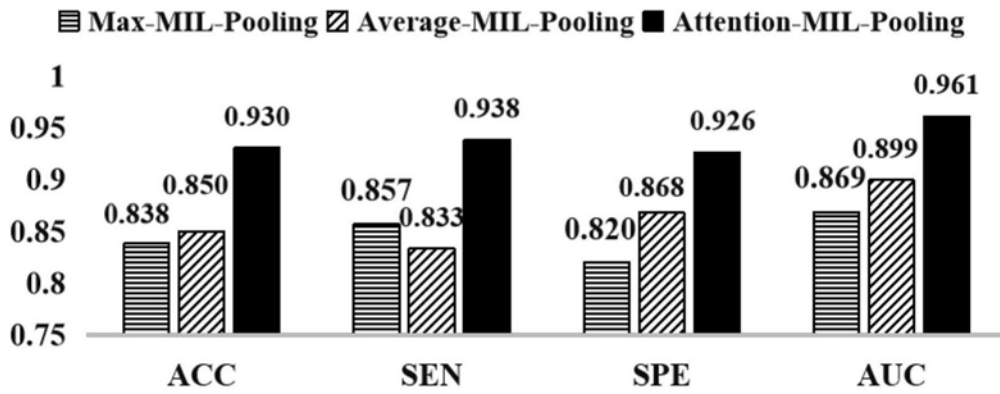


图2

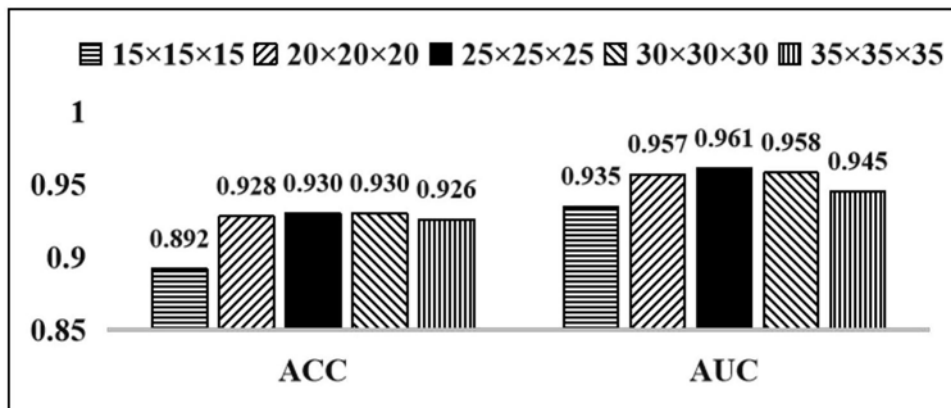


图3

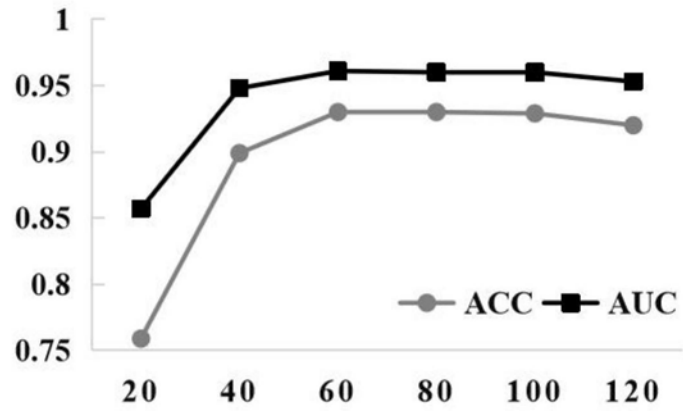


图4