



(12)发明专利申请

(10)申请公布号 CN 106780482 A

(43)申请公布日 2017.05.31

(21)申请号 201710011904.8

(22)申请日 2017.01.08

(71)申请人 广东工业大学

地址 510062 广东省广州市越秀区东风东  
路729号

(72)发明人 张钢 林逸禹 赖华东 林栋

(74)专利代理机构 广东广信君达律师事务所  
44329

代理人 杨晓松

(51)Int.Cl.

G06T 7/00(2017.01)

G06K 9/62(2006.01)

G06T 3/60(2006.01)

权利要求书2页 说明书6页

(54)发明名称

一种医学图像分类方法

(57)摘要

本发明提出一种基于图像特征迁移学习和深度学习的医学图像分类方法,通过已经训练好的在图像理解领域中表现良好的模型,对医学图像进行特征提取,然后把得到的特征图用于训练一个能够给出最后分类的卷积神经网络模型,本发明能够充分利用已有在图像理解领域表现良好的模型所提取的特征,在小规模的医学图像数据集上能够取得较好的效果。

1. 一种医学图像分类方法,其特征在于,所述方法具体包括以下步骤:原始医学图像的归一化处理、医学图像数据集的扩充、医学图像的特征图提取、卷积神经网络的设计、卷积神经网络的训练以及医学图像的分类。

2. 如权利要求1所述的方法,其特征在于,所述原始医学图像的归一化处理的步骤具体包括:

获取原始医学图像为RGB通道的彩色图像,图像大小为 $h*w$ ,整个图像是一个 $h*w*3$ 的矩阵,每个通道是一个 $h*w$ 的矩阵,对每个通道分别进行归一化:先求出在该幅图像中每个通道的像素最大值和最小值,分别记为 $p^{max}$ 和 $p^{min}$ ,对于该通道的每个像素值进行归一化:

$p^* = \frac{p - p^{min}}{p^{max} - p^{min}}$ ,其中 $p^*$ 为归一化之后的像素值, $p$ 为归一化之前的像素值。

3. 如权利要求2所述的方法,其特征在于,所述医学图像数据集的扩充的步骤具体包括:

-翻转:对经过归一化之后的原始图像进行水平翻转和垂直翻转,把翻转后的图像作为新的图像加入到数据集中,则数据集中图像的增加量为 $2n$ ;

-旋转:对经过归一化之后的原始图像进行90、180、270度的旋转,把旋转后的图像作为新的图像加入到数据集中,则数据集中图像的增加量为 $3n$ ;

-加入误差:对经过翻转和旋转之后的图像的每一个像素点加入随机误差,误差在(0, 0.005)之间均匀分布,把加入像素误差的图像加入到数据集中;

-挖空像素:随机把经过翻转和旋转之后的图像 $p\%$ 的像素挖空,所述挖空指把被选中的像素随机填上全白或全黑,把挖空像素的图像加入到数据集中;经过图像扩充操作后,图像数据集的规模变为 $25n$ ,其中, $n$ 为原始图像的张数。

4. 如权利要求3所述的方法,其特征在于,所述医学图像的特征图提取的步骤具体包括:

-把数据集中的每一幅医学图像输入到选定的深度学习模型,则模型根据自身训练好的权值和激活函数,对输入的图像进行逐层次的特征提取,在每一层均会形成一系列的特征图,且在模型的输出层会有一个输出;

-对每一幅图像,随机选取经过模型后的所有层中维数在 $28*28-128*128$ 之间的特征图,一共选取 $v$ 次,其中 $v$ 为奇数。

5. 如权利要求4所述的方法,其特征在于,所述卷积神经网络的设计的步骤具体包括:

本步骤训练 $v$ 个卷积神经网络模型,每一个模型使用前一步骤所提取出来的特征图作为输入图像的特征表达,并且本步骤中卷积神经网络由输入层、卷积层/池化层、全连接层和输出层组成:

A. 输入层

输入层接受由前一步骤所提取出来的特征图数据,其输入维数与特征图的维数和抽样后的通道数相同;

B. 卷积层/池化层

-卷积层:选用卷积核为 $5*5$ 的卷积层,每一层的通道数与输入层相同,一个卷积层之后跟一个非线性的激活层,激活层的激活函数如下:

• ReLU:  $y = \max(x, 0)$ ,其中 $x$ 为上一层的输出, $y$ 为本激活层的输出;

-池化层:采用 $5*5$ 的卷积核,缩小比率为2;池化层与卷积核为 $5*5$ 的卷积层和激活单元组合使用,这样的组合一共有 $n$ 个;

C.全连接层:在最后一个池化层组合的输出上连接1个全连接层,该全连接层的维数为最后一个池化层组合输出维数的2倍,全连接层之后紧跟一个激活层;

D.输出层:输出层与最后一个激活层全连接,输出的维数与医学图像的分类数量相等。

6.如权利要求5所述的方法,其特征在于,所述卷积神经网络的训练的步骤具体包括:

采用标准的误差反向传播学习算法并根据网络的输入与输出的差异对网络的权值进行调整:对于每一个输入网络的医学图像特征图矩阵,网络的输出是一个与分类数量相等的向量,向量的每一位的取值为区间 $[0, 1]$ 之间的实数值,所述向量表示网络把输入医学图像判断为属于该位所代表的类别的概率;在计算每一幅图像的训练误差时,采用先离散化输出再用0-1误差计算总的误差,即对于每一个输出概率,若它的值大于0.5,则离散化为1,反之离散化为0,其中1代表属于该类,0代表不属于该类,

若模型的判断与真实的类别一致,则误差为0,反之误差为1,然后根据误差进行从输出端到输入端权值反向调整;网络的权值采用 $[0, 1]$ 之间的随机数进行初始化,进行多轮训练,所有训练数据输入到网络中并且完成权值调整为一轮,直到网络的输出误差不再下降为止。

7.如权利要求6所述的方法,其特征在于,所述医学图像的分类的步骤具体包括:

给定一幅待分类的医学图像,先按所述原始医学图像的归一化处理的步骤、所述医学图像数据集的扩充的步骤和所述医学图像的特征图提取步骤后得到 $v$ 组特征图,把所述 $v$ 组特征图按照维数分别输入到训练好的 $v$ 个卷积神经网络中,得到 $v$ 个输出向量之后进行离散化,即若它的值大于0.5,则离散化为1,反之离散化为0,然后对每一位进行投票,若在某一位置上 $v$ 个输出中1的个数占多数,则判断该医学图像属于该类别,反之判断该图像不属于该类别。

## 一种医学图像分类方法

### 技术领域

[0001] 本发明所属的技术领域为医学图像分析领域,具体为一种医学图像分类方法。

### 背景技术

[0002] 医学图像在很多医科的诊断中都有重要的作用,例如组织病理图像对于确定患病组织所患疾病的类型和严重程度有十分重要的意义,很多恶性疾病的最终确诊都需要借助组织病理图像和活体检查。但由于医学图像的识别需要大量的专业知识和临床经验,而且受到读图者的主观因素影响较大,容易造成诊断结果的不稳定甚至反转,因此使用机器学习的方法,从历史的医学图像诊断数据中学习读图的经验,并转化为计算机的程序,对读图结果的稳定性和正确性都有重要意义。此外,医学图像的计算机辅助诊断有助于减轻医生的工作量,使他们能够更加专注于特殊病例的诊断。

[0003] 现有的医学图像计算机自动识别方法主要是使用机器学习模型和统计学习模型提取医学图像的特征表达与其所表示的疾病之间的关系,针对某种疾病的医学图像,设计或选用特定的特征提取方法,得到图像的特征向量或矩阵,然后设计或选用合适的分类器或回归器,对医学图像进行疾病分类或对严重程度打分。文献“Ozdemir,E.&Demir,C.G.A Hybrid Classification Model for Digital Pathology Using Structural and Statistical Pattern Recognition.IEEE Trans.Med.Imaging,2013,32,474-483”提出了一个用于病理学图像分类的基于结构和统计模式识别的混合模型,把病理图像转化为一个属性图和一系列的查询参考子图,然后据此提取能够用于分类医学图像的两类特征,最后训练目标分类器。该方法遵循着特征提取和分类器训练的传统思路,只能针对特定的疾病进行训练,而且需要较大的样本集。

[0004] 现有方法的不足之处是:

[0005] (1).需要设计适合特定类型医学图像的特征提取方式。不同的特征对于不同类型的医学图像的作用大不相同,但设计合适的特征需要领域专家花费很多时间,同时需要大量的领域经验。

[0006] (2).为了得到效果较好的模型,需要规模较大的历史诊断数据集训练模型。规模小的诊断数据集使得在建立医学图像分类模型的时候不能使用复杂的模型,因会为导致过度拟合,而简单的模型难以表达出医学图像与目标疾病/病理特性之间的关系;但另一方面,对于特定医科,或针对某种疾病的医学图像来说,要得到足够大的训练数据集是很难的,因为医学上每个样本的获得都是需要实际操作的,因而获取大数据集的代价十分高昂。

[0007] 基于此,本发明提出一种基于图像特征迁移学习和深度学习的医学图像分类方法,通过已经训练好的在图像理解领域中表现良好的模型,对医学图像进行特征提取,然后把得到的特征图用于训练一个能够给出最后分类的卷积神经网络模型,本发明能够充分利用已有在图像理解领域表现良好的模型所提取的特征,在小规模的医学图像数据集上能够取得较好的效果。

## 发明内容

[0008] 本发明旨在克服现有的医学图像分类方法的不足,包括现有方法需要设计适合特定类型医学图像的特征提取方式、现有方法需要较大规模的医学图像数据集等。本发明的特征包括原始医学图像的归一化处理、医学图像数据集的扩充、利用已有的在图像理解领域表现良好的模型进行初步特征提取、卷积神经网络的训练和医学图像的分类,每个过程包含若干个步骤,其特征分别描述如下:

[0009] (1).原始医学图像的归一化处理

[0010] 归一化处理的目的是使图像的矩阵的数值转化为0至1之间的实数,转化为与训练过程的权值属于同一区间的数值。原始医学图像为RGB通道的彩色图像,设图像大小为 $h*w$ ,整个图像是一个 $h*w*3$ 的矩阵,每个通道是一个 $h*w$ 的矩阵,对每个通道分别进行归一化,归一化的方法是先求出在该幅图像中每个通道的像素最大值和最小值,分别记为 $p^{\max}$ 和 $p^{\min}$ ,对于该通道的每个像素值,归一化方法为: $p^* = \frac{p-p^{\min}}{p^{\max}-p^{\min}}$ ,其中 $p^*$ 为归一化之后的像素值, $p$ 为归一化之前的像素值。此外规定,如果 $p^{\max}$ 和 $p^{\min}$ 均为0,则归一化后的像素值仍为0。

[0011] (2).医学图像数据集的扩充

[0012] 医学图像数据集的扩充的目的是增加数据集的规模,突出能够决定最终分类标记的图像特征,降低模型训练时过度拟合的可能性。

[0013] 假设原始图像有 $n$ 张。通过以下四种手段进行图像层次扩充:

[0014] -翻转:对经过归一化之后的原始图像进行水平翻转和垂直翻转,把翻转后的图像作为新的图像加入到数据集中,则数据集中图像的增加量为 $2n$ ;

[0015] -旋转:对经过归一化之后的原始图像进行90、180、270度的旋转,把旋转后的图像作为新的图像加入到数据集中,则数据集中图像的增加量为 $3n$ ;

[0016] -加入误差:对经过翻转和旋转之后的图像的每一个像素点加入随机误差,误差在(0,0.005)之间均匀分布,把加入像素误差的图像加入到数据集中;

[0017] -挖空像素:对经过翻转和旋转之后的图像进行像素挖空,随机把图中的 $p\%$ 的像素( $p$ 在5-15之间),挖空操作是把被选中的像素随机填上全白(255,255,255)或全黑(0,0,0),把挖空像素的图像加入到数据集中;经过图像扩充操作后,图像数据集的规模变为 $(2n+3n)*2+(2n+3n)=25n$ 。

[0018] (3).医学图像的特征图提取

[0019] 利用已有的在图像理解领域表现良好的模型进行医学图像的特征图提取。具体做法如下:

[0020] -选取一个在图像理解领域表现良好的模型(已经训练好),例如在近几年的ImageNet图像识别大赛(ILSVRC,<http://image-net.org/challenges/LSVRC/>)上排名靠前的深度学习模型,这些模型在日常的图像理解问题上有较好的表现,而且每年都有比较明显的进步;

[0021] -把数据集中的每一幅医学图像输入到选定的深度学习模型,则模型会根据自身训练好的权值和激活函数,对输入的图像进行逐层次的特征提取,在每一层均会形成一系列的特征图,且在模型的输出层会有一个输出;

[0022] -对每一幅图像,随机选取经过模型后的所有层中维数在 $28*28-128*128$ 之间的特

征图,一共选取 $v$ 次,其中 $v$ 为奇数,有以下问题需要说明:

[0023] ■每一层的特征图的通道数可能不相同,例如有的层为 $28*28*512$ ,有的为 $32*32*256$ ,有的为 $64*64*128$ ,这取决于选用的已经训练好的模型的设计;

[0024] ■设定一个通道数阈值 $T^{\max}$ ,选取通道数大于等于 $T^{\max}$ 的层的输出特征图,策略如下:如果某一层的特征图通道数小于 $T^{\max}$ ,查看是否有其它层与该层的特征图维数相同,把所有与该层维数相同的特征图全部放在一起,如果其数量仍小于 $T^{\max}$ ,则放弃该维数的特征图,否则从中随机选取 $T^{\max}$ 个特征图作为图像的特征图;如果某一层的特征图通道数大于 $T^{\max}$ ,则直接从中随机选取 $T^{\max}$ 个特征图,作为图像的特征图;

[0025] 至此,训练集中的每一幅图像通过一个已经训练好的模型,被表达为 $v$ 种特征图表达。

[0026] (4).卷积神经网络的设计和训练

[0027] 本步骤训练 $v$ 个卷积神经网络模型,每一个模型使用步骤(3)所提取出来的一种特征图作为输入图像的特征表达。本步骤中卷积神经网络的结构如下:

[0028] 卷积神经网络由输入层、卷积层/池化层、全连接层和输出层组成。

[0029] A.输入层

[0030] 输入层接受由步骤(3)所提取出来的特征图数据,其输入维数与特征图的维数和抽样后的通道数相同;

[0031] B.卷积层/池化层

[0032] -卷积层:先用卷积核为 $5*5$ 的卷积层,每一层的通道数与输入层相同。一个卷积层之后跟一个非线性的激活层,激活层的激活函数如下:

[0033] • ReLU:  $y = \max(x, 0)$ ,其中 $x$ 为上一层的输出, $y$ 为本激活层的输出;

[0034] -池化层:采用 $5*5$ 的卷积核,缩小比率为2,遇到不能整除的情况采用水平单边补0的方法,即 $[0 \ 1 \ 0 \ 0]$ 或 $[0 \ 0 \ 0 \ 1]$ ;池化层与卷积核为 $5*5$ 的卷积层和激活单元组合使用,这样的组合一共有 $n$ 个;

[0035] C.全连接层:在最后一个池化层组合的输出上连接1个全连接层,该全连接层的维数为最后一个池化层组合输出维数的2倍,全连接层之后紧跟一个激活层;

[0036] D.输出层:输出层与最后一个激活层全连接,输出的维数与医学图像类别数量相等。

[0037] (5).训练过程

[0038] 采用标准的误差反向传播学习算法对网络的权值进行调整,调整是根据网络的输入与输出的差异进行的。具体来说,对于每一个输入网络的医学图像特征图矩阵,网络的输出是一个与分类数量相等的向量,向量的每一位的取值为区间 $[0, 1]$ 之间的实数值,代表网络把输入医学图像判断为属于该位所代表的类别的概率。在计算每一幅图像的训练误差时,采用先离散化输出再用0-1误差计算总的误差,即对于每一个输出概率,若它的值大于0.5,则离散化为1,反之离散化为0,其中1代表属于该类,0代表不属于该类,若模型的判断与真实的类别一致,则误差为0,反之误差为1,然后根据误差进行从输出端到输入端权值反向调整。

[0039] 网络的权值采用 $[0, 1]$ 之间的随机数进行初始化,进行多轮训练,所有训练数据输入到网络中并且完成权值调整为一轮,直到网络的输出误差不再下降为止。

## [0040] (6).医学图像分类

[0041] 给定一幅待分类的医学图像,先按本发明的步骤(1)中所述方法进行图像归一化,然后把归一化后的待分类图像输入到与训练集所使用的图像理解领域表现良好的模型中,按步骤(3)的方法,从模型中得到 $v$ 组特征图,把这 $v$ 组特征图按照维数分别输入到步骤(4)中训练好的 $v$ 个卷积神经网络中,得到 $v$ 个输出向量(每一位均为 $[0,1]$ 之间的实值),之后进行离散化,即若它的值大于0.5,则离散化为1,反之离散化为0,然后对每一位进行投票,若在某一位上 $v$ 个输出中1的个数占多数,则判断该医学图像属于该类别,反之判断该图像不属于该类别。

## 具体实施例

[0042] 本发明在一个公开的医学图像数据集上进行测试,取得了较好的效果。下面给出一个实施例,以公开的医学图像数据集histologyDS (<http://www.informed.unal.edu.co/histologyDS>) 作为测试数据集,该数据集为人体组织的病理图像数据集,有2828张图像,分为4个类别,图像为jpg格式,颜色空间是RGB,大小为720\*480像素。

### [0043] (1) 数据归一化

[0044] 按照第5点发明内容的步骤(1)进行图像的归一化,归一化是每张图像每个通道分别进行的,找出每个通道(R或G或B)的最大像素和最小像素,按步骤(1)的方法进行归一化。

### [0045] (2) 数据集扩充

[0046] 按照发明内容的第(2)点对原医学图像数据集进行扩充:

[0047] -翻转:进行图像的水平 and 垂直翻转,翻转后图像数据的增加量为 $2828*2=5656$ ;

[0048] -旋转:对经过归一化之后的原始图像进行90、180、270度的旋转,把旋转后的图像作为新的图像加入到数据集中,则数据集中图像的增加量为 $2828*3=8484$ ;

[0049] -加入误差:对经过翻转和旋转之后的图像的每一个像素点加入随机误差,误差在 $(0,0.005)$ 之间均匀分布,把加入像素误差的图像加入到数据集中;

[0050] -挖空像素:对经过翻转和旋转之后的图像进行像素挖空,随机把图中的 $p\%$ 的像素( $p$ 在5-15之间),挖空操作是把被选中的像素随机填上全白(255,255,255)或全黑(0,0,0),把挖空像素的图像加入到数据集中;

[0051] 经过图像扩充操作后,图像数据集的规模变为 $(2*2828+3*2828)*2+(2*2828+3*2828)=25*2828=70700$ 。

### [0052] (3).医学图像的特征提取

[0053] 选取一个在图像理解领域表现良好的深度学习模型作为特征提取模型,在本例中选取ResNet-152-dag (<http://www.vlfeat.org/matconvnet/models/imagenet-resnet-152-dag.mat>) 作为特征提取模型,该模型在ILSVRC 2012图像理解挑战中在所有接受挑战的模型中排名第1 (<http://www.vlfeat.org/matconvnet/pretrained/>),把医学图像数据集中的每一幅图像输入到ResNet-152-dag模型,并在其中维数为 $28*28$ 至 $28*28-128*128$ 的特征图中进行 $v=5$ 次随机选择,选择时取最大的通道数为1024,最终选中第17、25、46、59、99层的特征图,其维数/通道数分别为 $56*56*256$ 、 $56*56*256$ 、 $28*28*512$ 、 $28*28*512$ 、 $28*28*512$ ,由于这5层的通道数都没有超出最大通道数1024,所以直接选取这些层的所有特征图作为输入医学图像的特征图。

[0054] (3) 卷积神经网络设计

[0055] 采用卷积神经网络深度学习模型,设计5个卷积神经网络模型对输入医学图像的 $v=5$ 个特征图进行训练,模型按发明内容的第(5)点进行设计,表1给出了面向第1、2组特征图表达的网络结构,表2给出了面向第3、4、5组特征图表达的网络结构。

[0056] 表1. 面向第1、2组特征图表达的网络结构

[0057]

序号	类型	说明
1	输入层	无
2	卷积层	5*5 的卷积核, 通道数 256, 输出维数: 52*52*256
3	激活层	Relu 激活函数
4	池化层	5*5 的卷积核, 通道数 256, 填充[0 1 0 1], 输出维数: 25*25*256
5	卷积层	5*5 的卷积核, 通道数 256, 输出维数: 21*21* 512
6	激活层	Relu 激活函数

[0058]

7	池化层	5*5 的卷积核, 通道数 512, 输出维数: 9*9*512
8	卷积层	5*5 的卷积核, 通道数 512, 输出维数: 5*5* 1024
9	激活层	Relu 激活函数
10	卷积层	5*5 的卷积核, 通道数 1024, 输出维数: 1*1*2048
11	全连接层	通道数 4096
12	输出层	Softmax 函数, 输出维数: 4

[0059] 表2. 面向第3、4、5组特征图表达的网络结构



[0060]

序号	类型	说明
1	输入层	无
2	卷积层	5*5 的卷积核, 通道数 512, 输出维数: 24*24*512
3	激活层	Relu 激活函数
4	池化层	5*5 的卷积核, 通道数 512, 填充[0 1 0 1], 输出维数: 11*11*512
5	卷积层	5*5 的卷积核, 通道数 512, 输出维数: 7*7* 1024
6	激活层	Relu 激活函数
7	卷积层	5*5 的卷积核, 通道数 1024, 输出维数: 3*3* 2048
8	激活层	Relu 激活函数
9	卷积层	3*3 的卷积核, 通道数 2048, 输出维数: 1*1*2048
10	全连接层	通道数 4096
11	输出层	Softmax 函数, 输出维数: 4

[0061] (3).网络训练

[0062] 在MatConvNet中通过配置文件实现表1和表2中的网络结构,把数据集做成Matlab数据文件.mat格式,然后使用MatConvNet提供的训练脚本cnn\_train.m进行训练。训练进行30轮,使用的学习率为前10轮0.05,11-20轮0.005,21-30轮0.0005。训练的损失函数采用zero-one loss。每个模型经过30轮训练后,系统会生成30个.mat文件,分别保存了各轮训练结束时模型的参数,这些.mat文件为各轮训练好的模型,可用于对未知多示例样本的分类。

[0063] (4).医学图像分类

[0064] 使用每个模型的第30轮训练结束时的网络模型作为分类模型,对于一个待分类的医学图像,按照第5点发明内容的步骤(1)进行图像的归一化;之后输入ResNet-152-dag模型,提取第17、25、46、59、99层的特征图,把这5组特征图输入训练好的5个模型,得到5个分类向量,对这5个分类向量进行离散化,即若它的某一位的值大于0.5,则离散化为1,反之离散化为0,然后对每一位进行投票,若在某一位上5个输出中1的个数占多数,则判断该医学图像属于该类别,反之判断该图像不属于该类别。