



(12) 发明专利申请

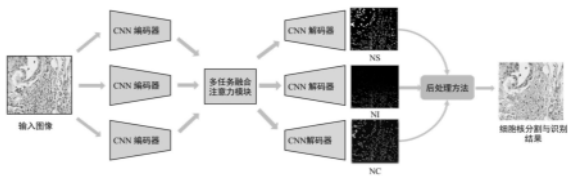
(10) 申请公布号 CN 114627142 A
(43) 申请公布日 2022. 06. 14

(21) 申请号 202210191562.3
(22) 申请日 2022.02.28
(71) 申请人 广东省人民医院
地址 510080 广东省广州市越秀区中山二
路106号
(72) 发明人 刘再毅 潘细朋 王瑛 程纪钧
韩楚 刘振丙 陆铖 陈浩 陈鑫
冯拯云 侯飞虎 梁长虹
(74) 专利代理机构 广州市华学知识产权代理有
限公司 44245
专利代理师 李斌
(51) Int.Cl.
G06T 7/12 (2017.01)
G06K 9/62 (2022.01)
G06V 10/764 (2022.01)

权利要求书3页 说明书9页 附图2页

(54) 发明名称
H&E染色组织病理图像细胞核分割与分类系
统、方法、设备及介质

(57) 摘要
本发明公开了一种H&E染色组织病理图像细胞核分割与分类系统、方法、设备及介质，系统包括多任务深度学习模块、多任务相关注意力机制模块、代价敏感学习策略模块、两步分水岭算法后处理模块；所述多任务深度学习模块，用于分别进行细胞核语义分割、实例分割和细胞核分类；所述多任务相关注意力机制模块，用于融合细胞核分割模块、细胞核实例分割模块和细胞核分类模块的特征表示，以利用全局上下文信息；所述代价敏感学习策略模块，用于有效解决不同细胞类型样本数量严重失衡导致模型性能下降的问题；所述两步分水岭算法后处理模块，用于对大小和形态上存在很大差异的细胞核进行更精确地分割。本发明可实现H&E染色组织病理学图像细胞核分割与分类任务。



1. H&E染色组织病理图像细胞核分割与分类系统,其特征在于,包括多任务深度学习模块、多任务相关注意力机制模块、代价敏感学习策略模块、两步分水岭算法后处理模块;

所述多任务深度学习模块包含三个子任务模块,所述三个子任务模块分别为细胞核分割模块、细胞核实例分割模块和细胞核分类模块,所述细胞核分割模块、细胞核实例分割模块和细胞核分类模块的输入图像分别通过编码结构、多任务相关注意力机制模块、解码结构后获得细胞核分割结果、实例分割结果、细胞核分类结果;

所述细胞核分割模块,用于区分输入图像的像素是否属于细胞核或非细胞核;

所述细胞核实例分割模块,用于生成实例分割的水平和垂直距离图,有助于对前一个模块的语义分割结果执行实例分割;

所述细胞核分类模块,用于获得组织病理图像对应的分类图,这有助于对前两个模块的实例分割结果完成细胞核分类;所述细胞核分类模块的预测结果对应细胞核实例中像素比最高的类别;

所述多任务相关注意力机制模块,用于融合细胞核分割模块、细胞核实例分割模块和细胞核分类模块的特征表示,实现全局上下文信息的利用;

所述代价敏感学习策略模块,用于有效解决不同细胞类型样本数量严重失衡导致模型性能下降的问题;

所述两步分水岭算法后处理模块,用于对大小、颜色和形态上存在很大差异的细胞核进行更精确地分割。

2. 根据权利要求1所述H&E染色组织病理图像细胞核分割与分类系统,其特征在于,所述多任务相关注意力机制模块通过基于多任务分支的特征交互机制,具体为:

三个子任务模块首先通过各自的编码器,然后连接到多任务相关注意力机制模块;

三个子任务模块向各自的解码器输出并执行各自的任务;其中,三个子任务模块都有各自的参数,而多任务相关注意力机制模块中的参数由所有子任务模块共享;由于多任务相关注意力机制模块的作用,每个子任务模块不仅学习自己的任务特征,同时学习到其他子任务模块的相关任务特征。

3. 根据权利要求1所述H&E染色组织病理图像细胞核分割与分类系统,其特征在于,所述代价敏感学习策略模块,通过在损失函数中嵌入代价敏感矩阵,减轻样本不平衡的影响,提高像素数量占比少的细胞核分类精度,具体为:

对于图像像素分类问题,将代价敏感的惩罚矩阵应用于每个patch中的每个像素,因此,首先使用现有的代价敏感矩阵W与输入样本图像中的每个像素相关联,输入尺寸为(h1, h2)的patch图像被定义为:

$$P = \begin{bmatrix} p(0,0), p(0,1), \dots, p(0,h2-1) \\ p(1,0), p(1,1), \dots, p(1,h2-1) \\ \dots & \dots & \dots & \dots \\ p(h1-1,0), p(h1-1,1), \dots, p(h1-1,h2-1) \end{bmatrix}$$

而 y_p 表示与P对应的真实值标签;

$$y_p = \begin{bmatrix} y(0,0), y(0,1), \dots, y(0,h2-1) \\ y(1,0), y(1,1), \dots, y(1,h2-1) \\ \dots \quad \dots \quad \dots \quad \dots \\ y(h1-1,0), y(h1-1,1), \dots, y(h1-1,h2-1) \end{bmatrix}$$

其中, $y_p(i, j) \in [0, M]$, 关联映射矩阵 M_p 被定义为:

$$M_p = \begin{bmatrix} w(y_{(0,0)},), w(y_{(0,1)},), \dots, w(y_{(0,h2-1)},) \\ w(y_{(1,0)},), w(y_{(1,1)},), \dots, w(y_{(1,h2-1)},) \\ \dots \quad \dots \quad \dots \quad \dots \\ w(y_{(h1-1,0)},), w(y_{(h1-1,1)},), \dots, w(y_{(h1-1,h2-1)},) \end{bmatrix}$$

代价敏感的损失函数定义为:

$$L^{cs}(\hat{y}, y) = L^{base}(\hat{y}, y) + \lambda(\hat{y}, M_p(y, \cdot))$$

\hat{y} 和 y 分别表示预测结果和真实标签, λ 表示代价敏感损失的权重, $M_p(y_i, \cdot)$ 表示真实标签与代价敏感标签相关联后的代价敏感混淆矩阵, L^{base} 和 $\lambda(\hat{y}, M_p(y, \cdot))$ 定义为:

$$L^{base} = L^{ce} = - \sum_{i=1}^N (y_{i_{one-hot}} * \log((\hat{y}_i)_{softmax}))$$

$$\lambda(\hat{y}, M_p(y, \cdot)) = \lambda * ((\hat{y}_i)_{softmax}) * M_p(y_i, \cdot)$$

因为 $w(k, k)$ 在代价敏感的惩罚矩阵 W 中为 0, $(\hat{y}_i)_{softmax}$ and $M_p(y_i, \cdot)$ 类别像素对应的操作结果为 0, 因此, 使用 L^{ce} 作为代价敏感损失函数的补充, 最后, L^{cs} 可用于类别不平衡的代价敏感惩罚。

4. 根据权利要求 1 所述 H&E 染色组织病理图像细胞核分割与分类系统, 其特征在于, 所述两步分水岭算法后处理模块, 包括下述两个步骤:

第一步分水岭算法: 采用梯度法获取水平方向和垂直方向的边缘特征, 并将两个方向得到的边缘特征叠加, 得到完整的边缘特征, 结合边缘结构特征, 引入标记控制分水岭方法, 获得细胞核实例分割结果;

第二步分水岭算法: 首先合并上述第一步分水岭算法错误划分的单元, 然后计算合并结果中单元的距离变换图, 在距离变换计算过程中, 距离背景越远的像素值越大, 结合距离变换图, 采用第二步标记控制分水岭算法, 对所有类型的单元实现更精确的实例分割。

5. 根据权利要求 4 所述 H&E 染色组织病理图像细胞核分割与分类系统, 其特征在于, 将细胞核分类分支应用于细胞核实例分割结果, 并根据类别属性对每个实例单元进行投票, 获得分割和分类结果。

6. 根据权利要求 1-5 中任一项所述 H&E 染色组织病理图像细胞核分割与分类系统的分类方法, 其特征在于, 包括下述步骤:

S1、获取 H&E 染色组织病理图像, 并将该病理图像输入多任务深度学习模块; 所述病理图像首先分别通过所述多任务深度学习模块对病理图像进行细胞语义分割、细胞核实例分割和细胞核分类的编码结构;

S2、多任务深度学习模块三个子任务的编码结构输出特征分别依次经过多任务相关注意力机制模块；

S3、多任务深度学习模块的三个子任务输出经过多任务相关注意力机制模块后，分别进入各自的解码结构，并分别获得细胞核分割、实例分割和细胞核分类的输出结果；

S4、多任务深度学习模块的细胞核分割结果和实例分割结果，通过第一次分水岭算法，获得第一次分水岭分割结果；

S5、多任务深度学习模块的细胞核分类结果与第一次分水岭分割结果进一步结合，通过统计第一步分水岭分割细胞中，所对应的细胞核分类结果，标识细胞核类型，获得细胞核分割与识别图；

S6、对于细胞核分割与识别图中被错误分割的细胞，现将相邻且类型形同的细胞合并，并计算合并结果的距离变换图，进行第二次标记控制分水岭算法，获得校正后的细胞核分割与识别结果。

7. 根据权利要求6所述H&E染色组织病理图像细胞核分割与分类系统的分类方法，其特征在于，所述第一步分水岭算法具体为：

采用梯度法获取水平方向和垂直方向的边缘特征，并将两个方向得到的边缘特征叠加，得到完整的边缘特征，结合边缘结构特征，引入标记控制分水岭方法，获得细胞核实例分割结果。

8. 根据权利要求7所述H&E染色组织病理图像细胞核分割与分类系统的分类方法，其特征在于，所述第二分水岭算法具体为：

首先合并上述第一步分水岭算法错误划分的单元，然后计算合并结果中单元的距离变换图，在距离变换计算过程中，距离背景越远的像素值越大，结合距离变换图，采用第二步标记控制分水岭算法，对所有类型的单元实现更精确的实例分割。

9. 一种电子设备，其特征在于，所述电子设备包括：

至少一个处理器；以及，

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的计算机程序指令，所述计算机程序指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行如权利要求1-5任一项所述的H&E染色组织病理图像细胞核分割与分类系统的分类方法。

10. 一种计算机可读存储介质，存储有程序，其特征在于，所述程序被处理器执行时，实现权利要求1-5任一项所述的H&E染色组织病理图像细胞核分割与分类系统的分类方法。

H&E染色组织病理图像细胞核分割与分类系统、方法、设备及介质

技术领域

[0001] 本发明属于数字图像处理的技术领域，具体涉及一种H&E染色组织病理图像细胞核分割与分类系统、方法、设备及介质。

背景技术

[0002] 癌症严重危害人类健康，病理诊断是判断肿瘤恶性程度的“金标准”。H&E染色是制备组织病理学切片的重要步骤之一，它可以显示肿瘤的组织结构和细胞成分。细胞核分割与分类是数字病理图像分析的必要前提和重要基础。现有的细胞核分割有基于区域的图像分割，基于区域的图像分割是依据自身的空间特征被划分为多个区域。常见的基于区域的图像分割方法有许多，譬如分水岭法和分裂合并法等等。基于区域的图像分割方法中，区域内的像素具有类似的梯度特征、纹理特征、像素灰度值特征等等。该类方法中目标像素的选择可以是具有众多相似特征的多个像素的小区域，或者可以是单个像素。如果所选择的目标像素的特征不一样，找到的像素也不一样，不一样的特征会影响到区域生长法最终的实现。当图像的像素的灰度值均匀分布且差异不大时，区域生长法的分割效果较好。但是，这个方法存在着一定的弊端：首先，其分割时需要自己预先设置相似特征，这将导致图像分割时具一定程度上的不确定性；其次，该方式对噪声相当敏感；还有，当图像的纹理信息比较复杂而且像素点较多时，该方法的分割速率会大幅度下降，图像分割的精确度也会受到干扰。在临床诊断中，较小比例的细胞同样发挥着非常重要的作用。因此，准确地分割和分类各个类别的细胞核，特别是样本图像中比例较小的细胞类别，具有重要临床意义和科学价值。

发明内容

[0003] 本发明的主要目的在于克服现有技术的缺点与不足，提供一种H&E染色组织病理图像细胞核分割与分类系统、方法、设备及介质，本发明由三个独立的编码器和解码器组成的多任务特征融合框架。进一步引入代价敏感学习策略，有效解决学习样本中不同细胞类型样本数量不均衡的问题。

[0004] 为了达到上述目的，本发明采用以下技术方案：

[0005] 本发明一方面提供了一种H&E染色组织病理图像细胞核分割与分类系统，包括多任务深度学习模块、多任务相关注意力机制模块、代价敏感学习策略模块、两步分水岭算法后处理模块；

[0006] 所述多任务深度学习模块包含三个子任务模块，所述三个子任务模块分别为细胞核分割模块、细胞核实例分割模块和细胞核分类模块，所述细胞核分割模块、细胞核实例分割模块和细胞核分类模块的输入图像分别通过编码结构、多任务相关注意力机制模块、解码结构后获得细胞核分割结果、实例分割结果、细胞核分类结果；

[0007] 所述细胞核分割模块，用于区分输入图像的像素是否属于细胞核或非细胞核；

[0008] 所述细胞核实例分割模块,用于生成实例分割的水平和垂直距离图,有助于对前一个模块的语义分割结果执行实例分割;

[0009] 所述细胞核分类模块,用于获得组织病理图像对应的分类图,这有助于对前两个模块的实例分割结果完成细胞核分类;所述细胞核分类模块的预测结果对应细胞核实例中像素比最高的类别;

[0010] 所述多任务相关注意力机制模块,用于融合细胞核分割模块、细胞核实例分割模块和细胞核分类模块的特征表示,实现全局上下文信息的利用;

[0011] 所述代价敏感学习策略模块,用于有效解决不同细胞类型样本数量严重失衡导致模型性能下降的问题;

[0012] 所述两步分水岭算法后处理模块,用于对大小、颜色和形态上存在很大差异的细胞核进行更精确地分割。

[0013] 作为优选的技术方案,所述多任务相关注意力机制模块通过基于多任务分支的特征交互机制,具体为:

[0014] 三个子任务模块首先通过各自的编码器,然后连接到多任务相关注意力机制模块;

[0015] 三个子任务模块向各自的解码器输出并执行各自的任务;其中,三个子任务模块都有各自的参数,而多任务相关注意力机制模块中的参数由所有子任务模块共享;由于多任务相关注意力机制模块的作用,每个子任务模块不仅学习自己的任务特征,同时学习到其他子任务模块的相关任务特征。

[0016] 作为优选的技术方案,所述代价敏感学习策略模块,通过在损失函数中嵌入代价敏感矩阵,减轻样本不平衡的影响,提高像素数量占比少的细胞核分类精度,具体为:

[0017] 对于图像像素分类问题,将代价敏感的惩罚矩阵应用于每个patch中的每个像素,因此,首先使用现有的代价敏感矩阵 W 与输入样本图像中的每个像素相关联,输入尺寸为 $(h1, h2)$ 的 patch图像被定义为:

$$[0018] \quad P = \begin{bmatrix} p(0,0), p(0,1), \dots, p(0,h2-1) \\ p(1,0), p(1,1), \dots, p(1,h2-1) \\ \dots \quad \dots \quad \dots \quad \dots \\ p(h1-1,0), p(h1-1,1), \dots, p(h1-1,h2-1) \end{bmatrix}$$

[0019] 而 y_p 表示与 P 对应的真实值标签;

$$[0020] \quad y_p = \begin{bmatrix} y(0,0), y(0,1), \dots, y(0,h2-1) \\ y(1,0), y(1,1), \dots, y(1,h2-1) \\ \dots \quad \dots \quad \dots \quad \dots \\ y(h1-1,0), y(h1-1,1), \dots, y(h1-1,h2-1) \end{bmatrix}$$

[0021] 其中, $y_p(i, j) \in [0, M]$,关联映射矩阵 M_p 被定义为:

$$[0022] \quad M_p = \begin{bmatrix} w(y_{(0,0)}), w(y_{(0,1)}), \dots, w(y_{(0,h2-1)}) \\ w(y_{(1,0)}), w(y_{(1,1)}), \dots, w(y_{(1,h2-1)}) \\ \dots \quad \dots \quad \dots \quad \dots \\ w(y_{(h1-1,0)}), w(y_{(h1-1,1)}), \dots, w(y_{(h1-1,h2-1)}) \end{bmatrix}$$

[0023] 代价敏感的损失函数定义为：

$$[0024] \quad L^{cs}(\hat{y}, y) = L^{base}(\hat{y}, y) + \lambda (\hat{y}, \mathbf{M}_p(y, \cdot))$$

[0025] \hat{y} 和 y 分别表示预测结果和真实标签, λ 表示代价敏感损失的权重, $\mathbf{M}_p(y_i, \cdot)$ 表示真实标签与代价敏感标签相关联后的代价敏感混淆矩阵, L^{base} 和 $\lambda (\hat{y}, \mathbf{M}_p(y, \cdot))$ 定义为：

$$[0026] \quad L^{base} = L^{ce} = - \sum_{i=1}^N (y_{i_{one-hot}} * \log((\hat{y}_i)_{softmax}))$$

$$[0027] \quad \lambda (\hat{y}, \mathbf{M}_p(y, \cdot)) = \lambda * ((\hat{y}_i)_{softmax}) * \mathbf{M}_p(y_i, \cdot)$$

[0028] 因为 $w(k, k)$ 在代价敏感的惩罚矩阵 W 中为0, $(\hat{y}_i)_{softmax}$ and $\mathbf{M}_p(y_i, \cdot)$ 类别像素对应的操作结果为0, 因此, 使用 L^{ce} 作为代价敏感损失函数的补充, 最后, L^{cs} 可用于类别不平衡的代价敏感惩罚。

[0029] 作为优选的技术方案, 所述两步分水岭算法后处理模块, 包括下述两个步骤：

[0030] 第一步分水岭算法：采用梯度法获取水平方向和垂直方向的边缘特征, 并将两个方向得到的边缘特征叠加, 得到完整的边缘特征, 结合边缘结构特征, 引入标记控制分水岭方法, 获得细胞核实例分割结果；

[0031] 第二步分水岭算法：首先合并上述第一步分水岭算法错误划分的单元, 然后计算合并结果中单元的距离变换图, 在距离变换计算过程中, 距离背景越远的像素值越大, 结合距离变换图, 采用第二步标记控制分水岭算法, 对所有类型的单元实现更精确的实例分割；

[0032] 将细胞核分类分支应用于细胞核实例分割结果, 并根据类别属性对每个实例单元进行投票, 获得分割和分类结果。

[0033] 本发明另一方面提供了一种H&E染色组织病理图像细胞核分割与分类系统的分类方法, 包括下述步骤：

[0034] S1、获取H&E染色组织病理图像, 并将该病理图像输入多任务深度学习模块；所述病理图像首先分别通过所述多任务深度学习模块对病理图像进行细胞语义分割、细胞核实例分割和细胞核分类的编码结构；

[0035] S2、多任务深度学习模块三个子任务的编码结构输出特征分别依次经过多任务相关注意力机制模块；

[0036] S3、多任务深度学习模块的三个子任务输出经过多任务相关注意力机制模块后, 分别进入各自的解码结构, 并分别获得细胞核分割、实例分割和细胞核分类的输出结果；

[0037] S4、多任务深度学习模块的细胞核分割结果和实例分割结果, 通过第一次分水岭算法, 获得第一次分水岭分割结果；

[0038] S5、多任务深度学习模块的细胞核分类结果与第一次分水岭分割结果进一步结合, 通过统计第一步分水岭分割细胞中, 所对应的细胞核分类结果, 标识细胞核类型, 获得细胞核分割与识别图；

[0039] S6、对于细胞核分割与识别图中被错误分割的细胞, 现将相邻且类型形同的细胞合并, 并计算合并结果的距离变换图, 进行第二次标记控制分水岭算法, 获得校正后的细胞核分割与识别结果。

[0040] 作为优选的技术方案,所述第一步分水岭算法具体为:

[0041] 采用梯度法获取水平方向和垂直方向的边缘特征,并将两个方向得到的边缘特征叠加,得到完整的边缘特征,结合边缘结构特征,引入标记控制分水岭方法,获得细胞核实例分割结果。

[0042] 作为优选的技术方案,所述第二分水岭算法具体为:

[0043] 首先合并上述第一步分水岭算法错误划分的单元,然后计算合并结果中单元的距离变换图,在距离变换计算过程中,距离背景越远的像素值越大,结合距离变换图,采用第二步标记控制分水岭算法,对所有类型的单元实现更精确的实例分割。

[0044] 本发明又一方面提供了一种电子设备,所述电子设备包括:

[0045] 至少一个处理器;以及,

[0046] 与所述至少一个处理器通信连接的存储器;其中,

[0047] 所述存储器存储有可被所述至少一个处理器执行的计算机程序指令,所述计算机程序指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行所述的H&E染色组织病理图像细胞核分割与分类系统的分类方法。

[0048] 本发明再一方面提供了一种计算机可读存储介质,存储有程序,所述程序被处理器执行时,实现所述的H&E染色组织病理图像细胞核分割与分类系统的分类方法。

[0049] 本发明与现有技术相比,具有如下优点和有益效果:

[0050] 本发明采用了多任务深度学习、多任务相关注意力机制、代价敏感学习策略、两步分水岭算法后处理的技术手段,解决了组织病理图像细胞核分割精度低、识别准确率低的技术问题,从而达到了精确分割和识别组织病理图像细胞核的技术效果。

附图说明

[0051] 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0052] 图1为本发明实施例H&E染色组织病理图像细胞核分割与分类系统的结构示意图。

[0053] 图2为本发明实施例H&E染色组织病理图像细胞核分割与分类方法的流程图。

[0054] 图3为本发明实施例中电子设备的结构示意图。

具体实施方式

[0055] 为了使本技术领域的人员更好地理解本申请方案,下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述。显然,所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0056] 在本申请中提及“实施例”意味着,结合实施例描述的特定特征、结构或特性可以包含在本申请的至少一个实施例中。在说明书中的各个位置出现该短语并不一定均是指相同的实施例,也不是与其它实施例互斥的独立的或备选的实施例。本领域技术人员显式地

和隐式地理解的是,本申请所描述的实施例可以与其它实施例相结合。

[0057] 如图1所示,本实施例H&E染色组织病理图像细胞核分割与分类系统,包括多任务深度学习模块、多任务相关注意力机制模块、代价敏感学习策略模块、两步分水岭算法后处理模块;该系统中,所述多任务深度学习模块由三个分支组成,用于分别进行细胞核语义分割、实例分割和细胞核分类;三个分支分别是细胞核分割(NS)分支、细胞核实例化(NI)分支和细胞核分类(NC)分支。这三个分支分别履行各自的职责,学习病理图像中核的形状、边缘和类别特征。NS分支区分输入图像的像素是否属于细胞核。该方法对组织病理图像中的细胞核进行了分割,得到了粗略的语义分割结果。然而,肿瘤的细胞核往往是簇聚的,导致大量的粘连和重叠细胞核。语义分割不能有效地分割细胞核簇。NI分支生成用于实例分割的水平和垂直距离图,有助于对前一个分支的语义分割结果执行实例分割。NC分支可以获得组织病理学图像对应的分类图,这有助于对前两个分支的实例分割结果完成细胞核分类。NC分支的预测结果为对应像素比最高的细胞核类别。

[0058] 为了有效地整合这些特征,本申请针对这三个分支分别在编码器和解码器结构之间引入了多任务相关注意策略。解码器完成后,NS分支、NI分支和NC分支分别进行细胞核语义分割、核实例化和细胞核分类。最后,提出了一个带有两步分水岭的后处理模块,以缓解细胞核的过分割现象。

[0059] 所述多任务相关注意力模块用于有效地整合三个独立的分支任务,有效地融合和学习不同子任务分支的特征表示,实现全局上下文信息的利用;

[0060] 所述代价敏感学习策略模块,用于有效解决不同细胞类型样本数量严重失衡导致模型性能下降的问题;

[0061] 所述两步分水岭算法后处理模块,能对大小、颜色和形态上存在很大差异的细胞核进行更精确地分割。

[0062] 进一步的,所述所述多任务深度学习模块细胞核分割模块、细胞核实例分割模块和细胞核分类模块;所述细胞核分割模块,用于区分输入图像的像素是否属于细胞核或非细胞核;

[0063] 所述细胞核实例分割模块,用于生成实例分割的水平和垂直距离图,有助于对前一个模块的语义分割结果执行实例分割;

[0064] 所述细胞核分类模块,用于获得组织病理图像对应的分类图,这有助于对前两个模块的实例分割结果完成细胞核分类,细胞核分类模块的预测结果对应细胞核实例中像素比最高的类别。

[0065] 可以理解的是,本发明在整个实验中所使用的数据样本来自于结直肠病理图像细胞核分割和表型 (CoNSeP) 和多器官细胞核分割和分类挑战赛 (MoNuSAC) 开源公开的数据,且实验均是回顾性实验,并不是在活体上所作的实验。

[0066] 进一步的,所述多任务相关注意力机制模块基于多任务分支的特征交互机制通过以下方式实现:

[0067] 首选,三个子任务分支首先通过各自的编码器,然后连接到注意力机制模块。

[0068] 之后,注意力机制模块向解码器输出并执行各自的任務。其中,三个子任务分支都有各自的参数,而注意力机制模块中的参数由所有任务分支共享。由于多任务相关注意力机制的作用,每个分支不仅学习自己的任务特征,而且还要考虑其他分支的特征。因此,与

常用的多任务学习方法相比,多任务相关注意力机制可以最大限度地减少不同子任务之间的干扰,同时融合各子任务的特征表示。注意力机制可以通过长期依赖性获取上下文信息,引入交叉注意力机制来构建多任务相关注意力机制模型,每个像素都可以在其交叉路径上获取像素的上下文信息。假设输入图像的宽度为 w ,高度为 h ,则每个交叉注意力模块的参数为 $w+h-1$ 。当连续执行两次操作时,每个像素将获取整个图像中所有像素的信息。因此,本发明将两个紧密的交叉注意力模块连接起来构建多任务相关注意力机制,既可以尽可能地减少计算量,又可以实现全局上下文信息的利用。

[0069] 进一步的,所述代价敏感学习策略模块,有效解决不同细胞核类型样本数量严重失衡的问题。该模块通过以下方式实现:

[0070] 数据不均衡是目标检测或分类任务中常见的问题,在组织病理图像数据集中,不同类型细胞数量的不平衡问题尤为突出。细胞类型的不平衡会导致分割和分类性能的下降,特别是数量少的细胞。为了解决这一问题,本申请提出了一种新的代价敏感损失函数,用于多类细胞核的分割和分类。通过在损失函数中嵌入代价敏感矩阵,可以减轻样本不平衡的影响,提高像素数量占比少的细胞核分类精度。

[0071] 对于多类任务,假设类别的数量为 M ,代价敏感矩阵 $W(M \times M)$ 可以定义如下:

$$[0072] \quad W = \begin{bmatrix} w(0,0), & w(0,1), & \dots, & w(0,M-1) \\ w(1,0), & w(1,1), & \dots, & w(1,M-1) \\ \dots & \dots & \dots & \dots \\ w(M-1,0), & w(M-1,1), & \dots, & w(M-1,M-1) \end{bmatrix}$$

[0073] $w(j,k)$ 为当类别 j 被错误地分类为类别 k 时的损失,矩阵 W 的对角元素是0,代价敏感矩阵上非对角线元素的值根据样本个数的比例进行设置。

[0074] 上述分析是基于图像级或patch级的。对于图像像素分类问题,本申请将代价敏感的惩罚矩阵应用于每个patch中的每个像素。因此,本申请首先使用现有的代价敏感矩阵 W 与输入样本图像中的每个像素相关联。输入尺寸为 $(h1,h2)$ 的patch图像可以被定义为:

$$[0075] \quad P = \begin{bmatrix} p(0,0), & p(0,1), & \dots, & p(0,h2-1) \\ p(1,0), & p(1,1), & \dots, & p(1,h2-1) \\ \dots & \dots & \dots & \dots \\ p(h1-1,0), & p(h1-1,1), & \dots, & p(h1-1,h2-1) \end{bmatrix}$$

[0076] 而 y_p 表示与 P 对应的真实值标签。

$$[0077] \quad y_p = \begin{bmatrix} y(0,0), & y(0,1), & \dots, & y(0,h2-1) \\ y(1,0), & y(1,1), & \dots, & y(1,h2-1) \\ \dots & \dots & \dots & \dots \\ y(h1-1,0), & y(h1-1,1), & \dots, & y(h1-1,h2-1) \end{bmatrix}$$

[0078] 其中, $y_p(i,j) \in [0,M]$ 。关联映射矩阵 M_p 被定义为:

$$[0079] \quad M_p = \begin{bmatrix} w(y_{(0,0)},), & w(y_{(0,1)},), & \cdots, & w(y_{(0,h2-1)},) \\ w(y_{(1,0)},), & w(y_{(1,1)},), & \cdots, & w(y_{(1,h2-1)},) \\ \cdots & \cdots & \cdots & \cdots \\ w(y_{(h1-1,0)},), & w(y_{(h1-1,1)},), & \cdots, & w(y_{(h1-1,h2-1)},) \end{bmatrix}$$

[0080] 代价敏感的损失函数可定义为:

$$[0081] \quad L^{cs}(\hat{y}, y) = L^{base}(\hat{y}, y) + \lambda (\hat{y}, M_p(y, \cdot))$$

[0082] \hat{y} 和 y 分别表示预测结果和真实标签, λ 表示代价敏感损失的权重, $M_p(y_i, \cdot)$ 表示真实标签与代价敏感标签相关联后的代价敏感混淆矩阵, L^{base} 和 $\lambda (\hat{y}, M_p(y, \cdot))$ 定义为:

$$[0083] \quad L^{base} = L^{ce} = - \sum_{i=1}^N (y_{i_{one-hot}} * \log((\hat{y}_i)_{softmax}))$$

$$[0084] \quad \lambda (\hat{y}, M_p(y, \cdot)) = \lambda * ((\hat{y}_i)_{softmax}) * M_p(y_i, \cdot)$$

[0085] 因为 $w(k, k)$ 在代价敏感的惩罚矩阵 W 中为0, $(\hat{y}_i)_{softmax}$ and $M_p(y_i, \cdot)$ 类别像素对应的操作结果为0, 因此, 本申请使用 L^{ce} 作为代价敏感损失函数的补充。最后, L^{cs} 可用于类别不平衡的代价敏感惩罚。

[0086] 进一步的, 所述两步分水岭算法后处理模块, 可以提高细胞核分割和分类的性能。该模块通过以下方式实现:

[0087] 解码器阶段结束后, 细胞核实例分割分支得到基于细胞核实例的水平和垂直距离图。本申请采用梯度法获取水平方向和垂直方向的边缘特征, 并将两个方向得到的边缘特征叠加, 得到完整的边缘特征。结合边缘结构特征, 引入标记控制分水岭方法, 获得细胞核实例分割结果。然而, 仅仅基于水平和垂直距离的标记控制的分水岭方法并不能实现对所有类型细胞的精确实例分割, 因为不同类型的细胞在大小、颜色和形态上存在很大差异。例如, 巨噬细胞很容易被过度分割成多个实例。因此, 本申请采用一种基于距离变换的标记控制分水岭方法, 它将优化前一步的分割结果。在第二步中, 本申请合并错误划分的单元, 然后计算合并结果中单元的距离变换图。在距离变换计算过程中, 距离背景越远的像素值越大。结合距离变换图, 采用第二步标记控制分水岭算法, 对所有类型的单元实现更精确的实例分割。最后, 将细胞核分类分支应用于细胞核实例分割结果, 并根据类别属性对每个实例单元进行投票, 获得分割和分类结果。这样, 得到细胞核实例分割和分类结果。

[0088] 下面, 结合具体的实验对本发明的技术方案做进一步的说明:

[0089] 如图2所示, 本发明提供一种H&E染色组织病理图像细胞核分割与分类方法, 包括输入图像、编码器、多任务相关注意力、解码器、后处理和输出结果等几个步骤, 具体为:

[0090] 首先, 输入一个 256×256 大小的patch进入本申请的编码器模块, 编码器模块用于获取高级语义信息。编码器模块的第一至第四层对应于ResNet-34中的第一至第四残差块。每个残差模块包含批量归一化和修正线性单元激活函数。

[0091] 进一步地, 所述多任务融合注意力机制, 用于多任务分支的编解码器。三个子任务分支首先通过各自的编码器, 然后连接到注意力机制模块。由于多任务融合注意力的集成

机制,每个分支不仅学习自己的任务特征,而且还要考虑其他分支的特征。

[0092] 进一步地,它们向解码器输出并执行各自的任務。三个子任务都有各自的参数,而注意力模块中的参数由所有任务分支共享。

[0093] 进一步地,解码器完成后,NS分支、NI分支和NC分支分别进行细胞核语义分割、核实例化和细胞核分类。

[0094] 进一步地,所述两步分水岭算法后处理模块,分两步进行:

[0095] 第一步,本申请采用一种基于距离变换的标记控制分水岭方法,如图2所示,它将优化前一步的分割结果。

[0096] 第二步,本申请合并错误划分的单元,然后计算合并结果中单元的距离变换图。在距离变换计算过程中,距离背景越远的像素值越大。结合距离变换图,采用第二步标记控制分水岭算法,对所有类型的单元实现更精确的实例分割。最后,将细胞核分类分支应用于细胞核实例分割结果,并根据类别属性对每个实例单元进行投票,获得分割和分类结果。

[0097] 请参阅图3,在一个实施例中,提供了一种H&E染色组织病理图像细胞核分割与分类的方法的电子設備,所述电子设备300可以包括第一处理器301、第一存储器302和总线,还可以包括存储在所述第一存储器302中并可在所述第一处理器301上运行的计算机程序,如H&E染色组织病理图像细胞核分割与分类程序303。

[0098] 其中,所述第一存储器302至少包括一种类型的可读存储介质,所述可读存储介质包括闪存、移动硬盘、多媒体卡、卡型存储器(例如:SD或DX存储器等)、磁性存储器、磁盘、光盘等。所述第一存储器302在一些实施例中可以是电子设备300的内部存储单元,例如该电子设备300的移动硬盘。所述第一存储器302在另一些实施例中也可以是电子设备300的外部存储设备,例如电子设备300上配备的插接式移动硬盘、智能存储卡(Smart Media Card, SMC)、安全数字(SecureDigital,SD)卡、闪存卡(Flash Card)等。进一步地,所述第一存储器302还可以既包括电子设备300的内部存储单元也包括外部存储设备。所述第一存储器302不仅可以用于存储安装于电子设备300的应用软件及各类数据,例如H&E染色组织病理图像细胞核分割与分类程序303的代码等,还可以用于暂时地存储已经输出或者将要输出的数据。

[0099] 所述第一处理器301在一些实施例中可以由集成电路组成,例如可以由单个封装的集成电路所组成,也可以是由多个相同功能或不同功能封装的集成电路所组成,包括一个或者多个中央处理器(Central Processing unit,CPU)、微处理器、数字处理芯片、图形处理器及各种控制芯片的组合等。所述第一处理器301是所述电子设备的控制核心(Control Unit),利用各种接口和线路连接整个电子设备的各个部件,通过运行或执行存储在所述第一存储器302内的程序或者模块(例如联邦学习防御程序等),以及调用存储在所述第一存储器302内的数据,以执行电子设备300的各种功能和处理数据。

[0100] 图3仅示出了具有部件的电子設備,本领域技术人员可以理解的是,图3示出的结构并不构成对所述电子设备300的限定,可以包括比图示更少或者更多的部件,或者组合某些部件,或者不同的部件布置。

[0101] 所述电子设备300中的所述第一存储器302存储的H&E染色组织病理图像细胞核分割与分类程序303是多个指令的组合,在所述第一处理器301中运行时,可以实现:

[0102] S1、获取H&E染色组织病理图像,并将该病理图像输入多任务深度学习模块;所述

病理图像首先分别通过所述多任务深度学习模块对病理图像进行细胞语义分割、细胞核实例分隔和细胞核分类的编码结构；

[0103] S2、多任务深度学习模块三个子任务的编码结构输出特征分别依次经过多任务相关注意力机制模块；

[0104] S3、多任务深度学习模块的三个子任务输出经过多任务相关注意力机制模块后，分别进入各自的解码结构，并分别获得细胞核分割、实例分割和细胞核分类的输出结果；

[0105] S4、多任务深度学习模块的细胞核分割结果和实例分割结果，通过第一次分水岭算法，获得第一次分水岭分割结果；

[0106] S5、多任务深度学习模块的细胞核分类结果与第一次分水岭分割结果进一步结合，通过统计第一步分水岭分割细胞中，所对应的细胞核分类结果，标识细胞核类型，获得细胞核分割与识别图；

[0107] S6、对于细胞核分割与识别图中被错误分割的细胞，现将相邻且类型形同的细胞合并，并计算合并结果的距离变换图，进行第二次标记控制分水岭算法，获得校正后的细胞核分割与识别结果。

[0108] 进一步地，所述电子设备300集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用，可以存储在一个非易失性计算机可读取存储介质中。所述计算机可读介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM, Read-Only Memory)。

[0109] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，所述的程序可存储于一非易失性计算机可读取存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器(ROM)、可编程ROM(PROM)、电可编程ROM(EPROM)、电可擦除可编程ROM(EEPROM)或闪存。易失性存储器可包括随机存取存储器(RAM)或者外部高速缓冲存储器。作为说明而非局限，RAM以多种形式可得，诸如静态RAM(SRAM)、动态RAM(DRAM)、同步DRAM(SDRAM)、双数据率SDRAM(DDRSDRAM)、增强型SDRAM(ESDRAM)、同步链路(Synchlink)DRAM(SLDRAM)、存储器总线(Rambus)直接RAM(RDRAM)、直接存储器总线动态RAM(DRDRAM)、以及存储器总线动态RAM(RDRAM)等。

[0110] 以上实施例的各技术特征可以进行任意的组合，为使描述简洁，未对上述实施例中的各个技术特征所有可能的组合都进行描述，然而，只要这些技术特征的组合不存在矛盾，都应当认为是本说明书记载的范围。

[0111] 上述实施例为本发明较佳的实施方式，但本发明的实施方式并不受上述实施例的限制，其他的任何未背离本发明的精神实质与原理下所作的改变、修饰、替代、组合、简化，均应为等效的置换方式，都包含在本发明的保护范围之内。

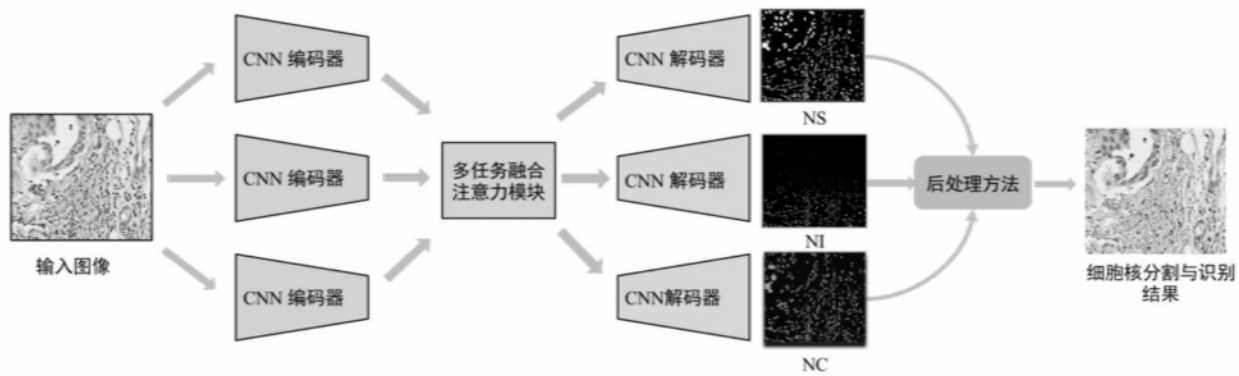


图1

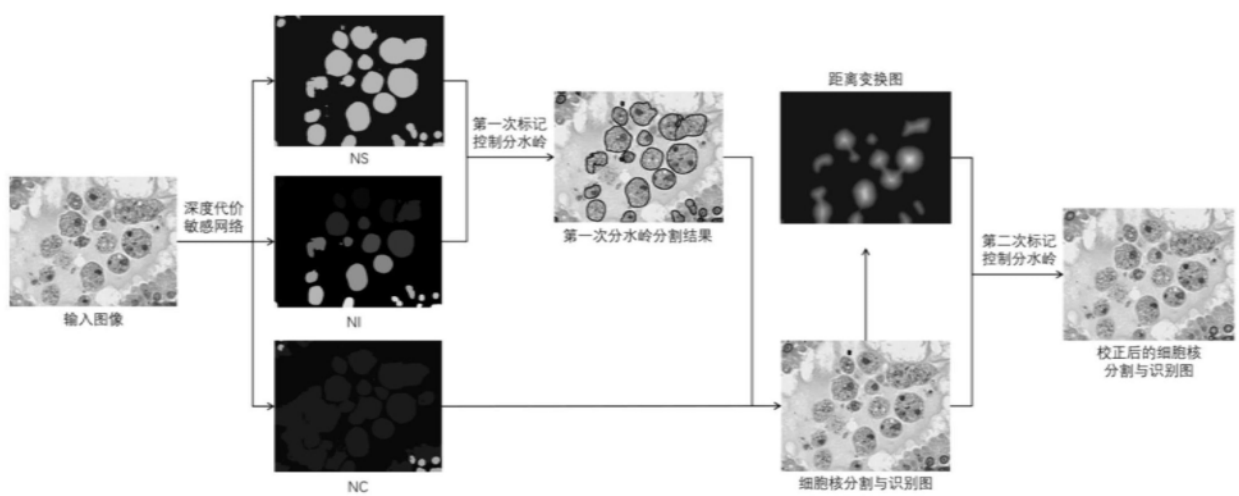


图2

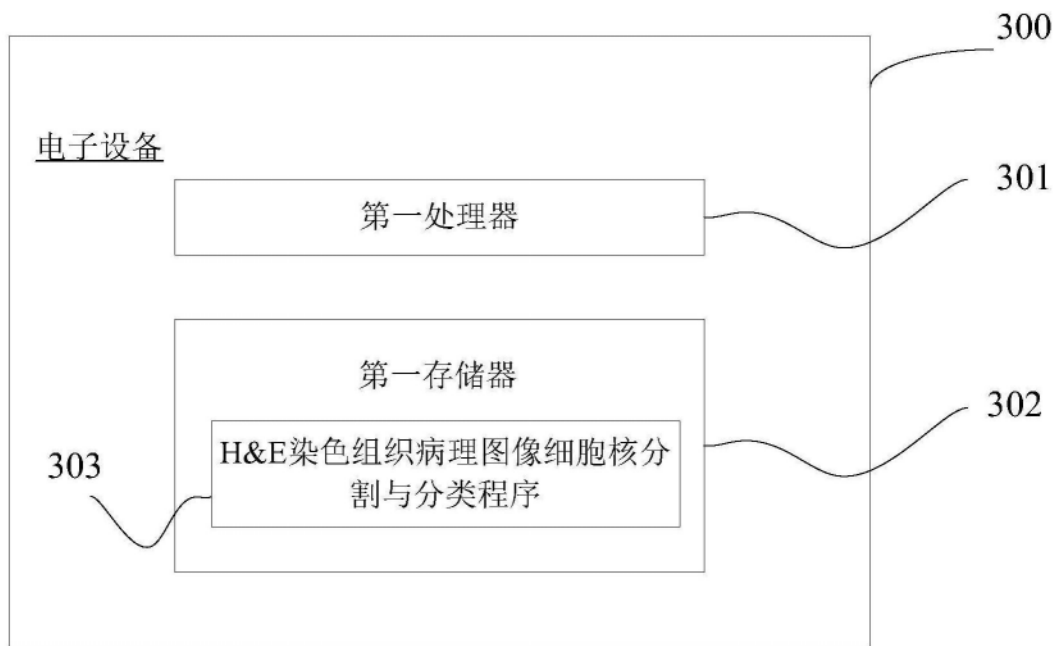


图3