



(12) 发明专利申请

(10) 申请公布号 CN 115564997 A

(43) 申请公布日 2023. 01. 03

(21) 申请号 202211199158.7

G06V 10/82 (2022.01)

(22) 申请日 2022.09.29

G16H 30/20 (2018.01)

(71) 申请人 东北大学

G06N 3/04 (2006.01)

地址 110004 辽宁省沈阳市和平区文化路3号巷11号

G06N 3/08 (2006.01)

申请人 中国医科大学附属第一医院

(72) 发明人 崔笑宇 高兴华 赵越 郑松

齐瑞群 郑亭亭 全浩 吴泽

宁爽迪 柏群 陈洪铎

(74) 专利代理机构 沈阳亚泰专利商标代理有限公司

公司 21107

专利代理师 王荣亮

(51) Int.Cl.

G06V 10/764 (2022.01)

G06V 10/80 (2022.01)

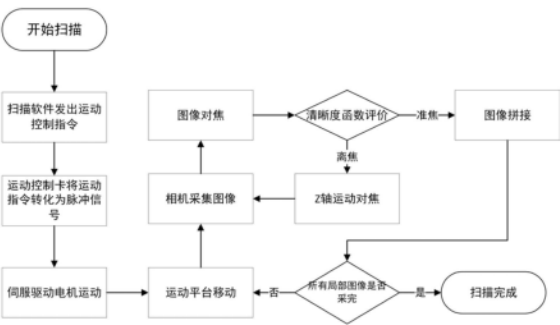
权利要求书4页 说明书9页 附图2页

(54) 发明名称

一种基于强化学习的病理切片扫描分析一体化方法及系统

(57) 摘要

一种基于强化学习的病理切片扫描分析一体化方法及系统,解决目前所使用的传统方法训练出的网络虽然能够得到一些较好的效果,但组织玻片的扫描过程和分析过程是依次单独进行的,人力物力消耗大,计算效率低的问题。该基于强化学习的病理切片扫描分析一体化方法及系统通过模型模仿病理专家的诊断过程,首先在低分辨率整张幻灯片图像中采样可疑的疾病区域,然后切换到高倍率下进一步确认;并通过卷积神经网络和循环神经网络对选定的感兴趣区域进行处理,学习不同疾病区域的判别特征,并且预测下一个位置。该一体化方法及系统能够使组织玻片扫描过程与人工智能分析并行,可以简化人力物力,病理判断效率高,实用性强。



1. 一种基于强化学习的病理切片扫描分析一体化方法,其特征在于,包括如下步骤:

步骤一、从皮肤病理数据库中,通过病理数字切片扫描仪收集数据,病理专家对数据进行进一步的确认,以保证数据标签的准确性;

步骤二、将收集的病理数据集,经过数据预处理,输入到基于卷积神经网络搭建的多实例专家指导模型,并采用交叉熵损失函数对专家指导模型进行训练;进而经过不断的训练,来获得一个能够指导深度强化学习的智能体进行决策的专家指导模型;

步骤三、将收集的病理数据集,分为训练集、验证集和测试集;当训练深度强化学习模型时,对训练集中的图像依次进行处理,设置迭代次数,限制每一张图像被迭代的次数;利用贪婪策略和动作网络获取智能体的动作,进而生成训练样本,并且,输入到基于多实例学习构建的专家模型和强化学习模型中;

步骤四、专家模型对输入的样本进行预测,其预测结果指导对强化学习智能体进行训练;采用策略梯度和奖励函数对智能体输出动作的网络进行训练,通过知识蒸馏的方式对智能体的决策分类网络进行训练;训练后的智能体生成下一个动作,作为新的样本输入到专家模型和强化学习模型,直到满足预先设置的终止条件,该图像的处理过程结束;

步骤五、返回步骤二,继续从训练集中输入下一个图像,直到训练集中的所有图像处理结束,强化学习网络参数确定;

步骤六、通过训练好的深度强化学习模型,对测试集中的图像进行疾病检测和分类,输出分类结果。

2. 根据权利要求1所述的基于强化学习的病理切片扫描分析一体化方法,其特征在于:所述步骤二,所建立的多实例专家指导模型如公式(1)所示:

$$y = f_{MILTM_{Res}}^{\theta}(x) \quad (1)$$

其中,x是经过专家视点约束的病理图像数据, $f_{MILTM_{Res}}^{\theta}$ 是具有可训练参数 θ 的多实例专家指导模型,使用ResNet卷积网络实现,y是卷积神经网络输出的预测结果;

用于指导模型训练的交叉熵损失函数如公式(2)所示:

$$L = \frac{1}{N} \sum_{i=1}^N L_i = \frac{1}{N} \sum_{i=1}^N -[Y_i \cdot \log(y_i) + (1 - Y_i) \cdot \log(1 - y_i)] \quad (2)$$

其中, Y_i 是第i张切片的标签, y_i 是模型输出的第i张切片的预测结果。

3. 根据权利要求1所述的基于强化学习的病理切片扫描分析一体化方法,其特征在于:所述步骤三,动作网络被定义为搜索智能体,用于输出动作,具体来说:使用长短记忆网络与全连接层结合设计了搜索智能体—SeAgent;并通过优化参数 $\phi_{SeAgent}$ 使SeAgent快速查找可疑区域;SeAgent输入的是状态 s_{it} 的特征向量 $v_{MILTM_Res}^{s_{it}}$; $v_{MILTM_Res}^{s_{it}}$ 表示 s_{it} 进入到MILTM经过卷积层、RULE非线性激活层、池化层、ResNet-34的四个block块(F_{MILTM}^{Res})输出的特征;使用预训练好的MILTM提取状态特征向量,不仅缩短了SeAgent特征提取的时间,而且提高了SeAgent的精准性。

4. 根据权利要求3所述的基于强化学习的病理切片扫描分析一体化方法,其特征在于:利用公式(5)把 $v_{MILTM_Res}^{s_{it}}$ 经过LSTM($F_{SeAgent}^{LSTM}$)与FC($F_{SeAgent}^{FC1}$)的特征

$v_{SeAgent_LSTM}^{s_{it}}$ 、 $v_{SeAgent_FC1}^{s_{it}}$ 进行融合,使得SeAgent充分的考虑了当前状态特征与先前状态信息;将融合后的特征经过全连接层 $F_{SeAgent}^{FC2}$ 和softamx函数,最终输出SeAgent的下一步动作 $\pi_{\Phi_{SeAgent}}(s_{it}, a_{it})$ 公式 (6) ;

$$v_{SeAgent_LSTM}^{s_{it}} = F_{SeAgent}^{LSTM}(v_{MILTM_Res}^{s_{it}} | \varphi_{SeAgent_LSTM}) \quad (3)$$

$$v_{SeAgent_FC1}^{s_{it}} = F_{SeAgent}^{LSTM}(v_{MILTM_Res}^{s_{it}} | \varphi_{SeAgent_FC1}) \quad (4)$$

$$v_{fusion}^{s_{it}} = v_{SeAgent_LSTM}^{s_{it}} + v_{SeAgent_FC1}^{s_{it}} \quad (5)$$

$$\begin{aligned} \pi_{\Phi_{SeAgent}}(a_{it} | s_{it}) &= P_{\Phi_{SeAgent}}(a_{it} | s_{it}, \varphi_{SeAgent}) \\ &= softmax(F_{SeAgent}^{FC2}(v_{fusion}^{s_{it}} | \varphi_{SeAgent_FC2})) \end{aligned} \quad (6)。$$

5. 根据权利要求4所述的基于强化学习的病理切片扫描分析一体化方法,其特征在于:由于SeAgent不可微,使用策略梯度进行优化参数 $\varphi_{SeAgent}$,以学习最优搜索策略 $\pi_{\Phi_{SeAgent}}(a_{it} | s_{it})$ 实现 R_i 的期望最大化,定义目标函数为:

$$L_{SeAgent} = E\left[\sum_{t=1}^T \gamma^{t-1} r_{it}\right] = \sum_{t=1}^T P_{\Phi_{SeAgent}}(a_{it} | s_{it}, \varphi_{SeAgent}) \gamma^{t-1} r_{it} \quad (7)$$

SeAgent每完成一次任务,对搜索网络进行一次参数更新;目标函数 $L_{SeAgent}$ 求梯度的公式为:

$$\nabla_{\Phi_{SeAgent}} L_{SeAgent} = \sum_{t=1}^T \gamma^{t-1} r_{it} \nabla_{\Phi_{SeAgent}} \log(\pi_{\Phi_{SeAgent}}(a_{it} | s_{it})) \quad (8)。$$

6. 根据权利要求1所述的基于强化学习的病理切片扫描分析一体化方法,其特征在于:所述步骤四,为了训练智能体输出动作的SeAgent网络,根据任务特点设计奖励函数;奖励函数是驱动SeAgent选择最大奖励值动作的,使用公式 (9),鼓励SeAgent高效快速的完成疾病诊断任务;如果SeAgent选择的动作获取的状态 s_{it} 经过MILTM鉴别为肿瘤,则反馈给SeAgent的+2的奖励,否则给予一个-2的奖励;此外,如果 s_{it}' 也被MILTM鉴别为肿瘤,则将SeAgent的奖励扩大到5;奖励是在每个行动之后立即产生的;

$$r_{it} = \begin{cases} -2 & \text{if } \hat{y}_{MILTM}^1 = 0 \\ +2 & \text{if } \hat{y}_{MILTM}^1 = 1 \text{ and } \hat{y}_{MILTM}^2 = 0 \\ +5 & \text{if } \hat{y}_{MILTM}^1 = 1 \text{ and } \hat{y}_{MILTM}^2 = 1 \end{cases} \quad (9)$$

\hat{y}_{MILTM}^1 和 \hat{y}_{MILTM}^2 表示在状态 s_{it} 和 s_{it}' 时, F_{MILTM}^1 和 F_{MILTM}^2 预测的结果;其中:

$$\hat{y}_{MILTM}^1 = \operatorname{argmax}(F_{MILTM}^1(s_{it} | \theta_{MILTM}^1)) \quad (10)$$

$$\hat{y}_{MILTM}^2 = \operatorname{argmax}(F_{MILTM}^2(s_{it}' | \theta_{MILTM}^2)) \quad (11)$$

最终,通过公式 (12),计算获得最大的累积回报来指导SeAgent进行优化, γ^{t-1} 是 r_{it} 在t时刻奖励的权重因子;

$$R_i = \sum_{t=1}^T \gamma^{t-1} r_{it} \quad (12)。$$

7. 根据权利要求1所述的基于强化学习的病理切片扫描分析一体化方法,其特征在于:

所述步骤四,对于决策智能体DeAgent部分,使用一层全连接构建,具体来说:通过优化全连接层 $F_{DeAgent}$ 的参数 $\psi_{DeAgent}$,实现对状态 s_{it} 特征向量 $v_{MILTM_Res}^{s_{it}}$ 的精准分类;由于病理组织图像种包含大量的非肿瘤组织区域,在没有像素级注释的情况下,仅使用组织图像的全局标签和交叉熵损失函数很难实现对DeAgent的优化;此外,SeAgent每一步的状态 s_{it} ,获取的patch的标签不一定与全局标签保持一致;

并且,受教师-学生模型的启发,将交叉熵损失函数与知识蒸馏的方式进行结合指导DeAgent学习;把状态 s_{it} 输入到MILTM模型,将最后一层softmax函数之前输出的patch良性和恶性 $Z_{MILTM}^{s_{it}}$ 信息经过平滑之后作为软标签 $\hat{y}_{soft}^{s_{it}}$,将 s_{it} 对应的组织切片 X_i 的标签 Y_i 经过one-hot编码作为硬标签 $\hat{y}_{hard}^{s_{it}}$;其中,用Kullback-Leibler散度作为软损失衡量MILTM与DeAgent预测良性和恶性概率的差异,用交叉熵损失函数作为硬损失计算DeAgent的预测结果 $Z_{DeAgent}^{s_{it}}$ 与真实标签之间的误差;设置超参数 ϵ 平衡软标签和硬标签在总损失中的比例;具体公式为:

$$L_{DeAgent} = \epsilon * H^2 * KLdiv(\hat{y}_{soft}^{s_{it}}, \hat{y}_{DeAgent}^{s_{it}}) + (1 - \epsilon) * CE(\hat{y}_{hard}^{s_{it}}, \hat{y}_{DeAgent}^{s_{it}}) \quad (13)$$

$$\hat{y}_{soft}^{s_{it}} = \frac{\exp(\frac{Z_{MILTM}^{s_{it}, k}}{H})}{\sum_{k=1}^K \exp(\frac{Z_{MILTM}^{s_{it}, k}}{H})} \quad (14)$$

$$\hat{y}_{DeAgent}^{s_{it}} = \frac{\exp(\frac{Z_{DeAgent}^{s_{it}, k}}{H})}{\sum_{k=1}^K \exp(\frac{Z_{DeAgent}^{s_{it}, k}}{H})} \quad (15)$$

其中,K是类别,H是温度系数,用H来调节软标签 $\hat{y}_{soft}^{s_{it}}$ 中包含的信息。

8.一种基于强化学习的病理切片扫描分析一体化系统,其特征在于:包括数字病理自动扫描系统和人工智能辅助分析系统。

9.根据权利要求8所述的基于强化学习的病理切片扫描分析一体化系统,其特征在于:所述数字病理自动扫描系统包括图像获取模块、硬件控制模块、图像处理模块和数据管理模块;

所述图像获取模块由光源、显微镜、相机组成,用于采集切片清晰的局部显微图像;

所述硬件控制模块,可根据切片扫描的需求,控制运动平台到达指定位置;

所述图像处理模块用于完成局部显微图像的预处理,与上述硬件控制模块配合采集清晰的局部显微图像,将局部显微图像配准、融合成为全景数字切片;

所述数据管理模块分为参数存储和图像数据管理两部分;其中,参数存储部分用于记录运动平台的运动参数和扫描参数,图像管理部分用于保存和载入显微图像。

10.根据权利要求8所述的基于强化学习的病理切片扫描分析一体化系统,其特征在于:所述人工智能辅助分析系统,利用设计好的人工智能算法,对所述病理图像进行分析,包括图片采集传输模块、图像预处理模块、分析模块、可视化模块;

所述图像采集传输模块用于将镜头拍摄的图片传输给人工智能系统进行分析;

所述图像预处理模块用于将传入的图片切割成适合的小块,以便后续的基于人工智能的分析;

所述分析模块通过模型对输入图片的类别、是否存在异常区域、异常区域的位置等进行分析并输出;

所述可视化模块用于当整张病理图扫描并分析结束后输出原病理图及其上的异常区域的框取,并可将异常区域进行放大观察。

一种基于强化学习的病理切片扫描分析一体化方法及系统

技术领域

[0001] 本发明属于医疗技术领域,具体涉及一种能够使组织玻片扫描过程与人工智能分析并行,可以简化人力物力,病理判断效率高,实用性强的基于强化学习的病理切片扫描分析一体化方法及系统。

背景技术

[0002] 除了极少数的癌症以外,绝大多数癌症最终的确诊都必须依靠病理检查,然后结合临床以及其他实验室检查结果,并根据主观经验综合定性诊断;可以说诊断病理学是鉴别癌症的基础和金标准。然而,由于病理切片的种类繁多、尺寸大,一名病理医生想要熟练地读懂病理切片往往需要多年的积淀,而且,也无法完全保证诊断的绝对正确性。

[0003] 如今,人工智能技术已经发展到辅助病理诊断方面,各种技术层出不穷,但基本上都是在数字化全切片的基础上进行,且使用各种人工智能算法用于疾病预测、良恶性分类或分割等。在现有的基于人工智能技术的辅助病理诊断中,最简单的方法就是从全切片图像(Whole Slide Image,WSI)中用滑动窗口或者随机抽取的方式提取样本,并训练一个能识别WSI疾病区域的模型来预测未知的组织切片上是否存在疾病;这种提取patch的方式不可避免的会存在与预测疾病没有诊断相关性的无效区域。目前所使用的一些传统方法训练出的网络虽然能够得到一些较好的效果,但组织玻片的扫描过程和分析过程是依次单独进行的,人力物力消耗大,且需要花费很长的时间,进而导致计算效率较低。故有必要对现有的病理切片扫描和分析的方法和系统予以改进。

发明内容

[0004] 本发明就是针对上述问题,提供一种能够使组织玻片扫描过程与人工智能分析并行,可以简化人力物力,病理判断效率高,实用性强的基于强化学习的病理切片扫描分析一体化方法及系统。

[0005] 本发明所采用的技术方案是:该基于强化学习的病理切片扫描分析一体化方法包括如下步骤:

[0006] 步骤一、从皮肤病理数据库中,通过病理数字切片扫描仪收集数据,病理专家对数据进行进一步的确认,以保证数据标签的准确性;

[0007] 步骤二、将收集的病理数据集,经过数据预处理,输入到基于卷积神经网络搭建的多实例专家指导模型,并采用交叉熵损失函数对专家指导模型进行训练;进而经过不断的训练,来获得一个能够指导深度强化学习的智能体进行决策的专家指导模型;

[0008] 步骤三、将收集的病理数据集,分为训练集、验证集和测试集;当训练深度强化学习模型时,对训练集中的图像依次进行处理,设置迭代次数,限制每一张图像被迭代的次数;利用贪婪策略和动作网络获取智能体的动作,进而生成训练样本,并且,输入到基于多实例学习构建的专家模型和强化学习模型中;

[0009] 步骤四、专家模型对输入的样本进行预测,其预测结果指导对强化学习智能体进

行训练；采用策略梯度和奖励函数对智能体输出动作的网络进行训练，通过知识蒸馏的方式对智能体的决策分类网络进行训练；训练后的智能体生成下一个动作，作为新的样本输入到专家模型和强化学习模型，直到满足预先设置的终止条件，该图像的处理过程结束；

[0010] 步骤五、返回步骤二，继续从训练集中输入下一个图像，直到训练集中的所有图像处理结束，强化学习网络参数确定；

[0011] 步骤六、通过训练好的深度强化学习模型，对测试集中的图像进行疾病检测和分类，输出分类结果。

[0012] 所述步骤二，所建立的多实例专家指导模型如公式 (1) 所示：

$$[0013] \quad y = f_{MILTM_{Res}}^{\theta}(x) \quad (1)$$

[0014] 其中， x 是经过专家视点约束的病理图像数据， $f_{MILTM_{Res}}^{\theta}$ 是具有可训练参数 θ 的多实例专家指导模型，使用ResNet卷积网络实现， y 是卷积神经网络输出的预测结果；

[0015] 用于指导模型训练的交叉熵损失函数如公式 (2) 所示：

$$[0016] \quad L = \frac{1}{N} \sum_{i=1}^N L_i = \frac{1}{N} \sum_{i=1}^N -[Y_i \cdot \log(y_i) + (1 - Y_i) \cdot \log(1 - y_i)] \quad (2)$$

[0017] 其中， Y_i 是第 i 张切片的标签， y_i 是模型输出的第 i 张切片的预测结果。

[0018] 所述步骤三，动作网络被定义为搜索智能体，用于输出动作，具体来说：使用长短记忆网络与全连接层结合设计了搜索智能体—SeAgent；并通过优化参数 $\varphi_{SeAgent}$ 使SeAgent快速查找可疑区域；SeAgent输入的是状态 s_{it} 的特征向量 $v_{MILTM_Res}^{s_{it}}$ ； $v_{MILTM_Res}^{s_{it}}$ 表示 s_{it} 进入到MILTM经过卷积层、RULE非线性激活层、池化层、ResNet-34的四个block块(F_{MILTM}^{Res})输出的特征；使用预训练好的MILTM提取状态特征向量，不仅缩短了SeAgent特征提取的时间，而且提高了SeAgent的精准性。

[0019] 另外，利用公式 (5) 把 $v_{MILTM_Res}^{s_{it}}$ 经过 $LSTM(F_{SeAgent}^{LSTM})$ 与 $FC(F_{SeAgent}^{FC1})$ 的特征 $v_{SeAgent_LSTM}^{s_{it}}$ 、 $v_{SeAgent_FC1}^{s_{it}}$ 进行融合，使得SeAgent充分的考虑了当前状态特征与先前状态信息；将融合后的特征经过全连接层 $F_{SeAgent}^{FC2}$ 和softamx函数，最终输出SeAgent的下一步动作 $\pi_{\Phi_{SeAgent}}(s_{it}, a_{it})$ 公式 (6)；

$$[0020] \quad v_{SeAgent_LSTM}^{s_{it}} = F_{SeAgent}^{LSTM}(v_{MILTM_Res}^{s_{it}} | \varphi_{SeAgent_LSTM}) \quad (3)$$

$$[0021] \quad v_{SeAgent_FC1}^{s_{it}} = F_{SeAgent}^{LSTM}(v_{MILTM_Res}^{s_{it}} | \varphi_{SeAgent_FC1}) \quad (4)$$

$$[0022] \quad v_{fusion}^{s_{it}} = v_{SeAgent_LSTM}^{s_{it}} + v_{SeAgent_FC1}^{s_{it}} \quad (5)$$

$$[0023] \quad \begin{aligned} \pi_{\Phi_{SeAgent}}(a_{it} | s_{it}) &= P_{\Phi_{SeAgent}}(a_{it} | s_{it}, \varphi_{SeAgent}) \\ &= softmax(F_{SeAgent}^{FC2}(v_{fusion}^{s_{it}} | \varphi_{SeAgent_FC2})) \end{aligned} \quad (6)$$

[0024] 而且，由于SeAgent不可微，使用策略梯度进行优化参数 $\varphi_{SeAgent}$ ，以学习最优搜索策略 $\pi_{\Phi_{SeAgent}}(a_{it} | s_{it})$ 实现 R_i 的期望最大化，定义目标函数为：

$$[0025] \quad L_{SeAgent} = E \left[\sum_{t=1}^T \gamma^{t-1} r_{it} \right] = \sum_{t=1}^T P_{\Phi_{SeAgent}}(a_{it} | s_{it}, \varphi_{SeAgent}) \gamma^{t-1} r_{it} \quad (7)$$

[0026] SeAgent每完成一次任务,对搜索网络进行一次参数更新;目标函数 $L_{SeAgent}$ 求梯度的公式为:

$$[0027] \quad \nabla_{\Phi_{SeAgent}} L_{SeAgent} = \sum_{t=1}^T \gamma^{t-1} r_{it} \nabla_{\Phi_{SeAgent}} \log(\pi_{\Phi_{SeAgent}}(a_{it} | s_{it})) \quad (8)$$

[0028] 所述步骤四,为了训练智能体输出动作的SeAgent网络,根据任务特点设计奖励函数;奖励函数是驱动SeAgent选择最大奖励值动作的,使用公式(9),鼓励SeAgent高效快速的完成疾病诊断任务;如果SeAgent选择的动作获取的状态 s_{it} 经过MILTM鉴别为肿瘤,则反馈给SeAgent的+2的奖励,否则给予一个-2的奖励;此外,如果 s_{it}' 也被MILTM鉴别为肿瘤,则将SeAgent的奖励扩大到5;奖励是在每个行动之后立即产生的;

$$[0029] \quad r_{it} = \begin{cases} -2 & \text{if } \hat{y}_{MILTM}^1 = 0 \\ +2 & \text{if } \hat{y}_{MILTM}^1 = 1 \text{ and } \hat{y}_{MILTM}^2 = 0 \\ +5 & \text{if } \hat{y}_{MILTM}^1 = 1 \text{ and } \hat{y}_{MILTM}^2 = 1 \end{cases} \quad (9)$$

[0030] \hat{y}_{MILTM}^1 和 \hat{y}_{MILTM}^2 表示在状态 s_{it} 和 s_{it}' 时, F_{MILTM}^1 和 F_{MILTM}^2 预测的结果;其中:

$$[0031] \quad \hat{y}_{MILTM}^1 = \operatorname{argmax}(F_{MILTM}^1(s_{it} | \theta_{MILTM}^1)) \quad (10)$$

$$[0032] \quad \hat{y}_{MILTM}^2 = \operatorname{argmax}(F_{MILTM}^2(s_{it}' | \theta_{MILTM}^2)) \quad (11)$$

[0033] 最终,通过公式(12),计算获得最大的累积回报来指导SeAgent进行优化, γ^{t-1} 是 r_{it} 在t时刻奖励的权重因子;

$$[0034] \quad R_i = \sum_{t=1}^T \gamma^{t-1} r_{it} \quad (12)$$

[0035] 所述步骤四,对于决策智能体DeAgent部分,使用一层全连接构建,具体来说:通过优化全连接层 $F_{DeAgent}$ 的参数 $\psi_{DeAgent}$,实现对状态 s_{it} 特征向量 $v_{MILTM_Res}^{s_{it}}$ 的精准分类;由于病理组织图像种包含大量的非肿瘤组织区域,在没有像素级注释的情况下,仅使用组织图像的全局标签和交叉熵损失函数很难实现对DeAgent的优化;此外,SeAgent每一步的状态 s_{it} ,获取的patch的标签不一定与全局标签保持一致;

[0036] 并且,受教师-学生模型的启发,将交叉熵损失函数与知识蒸馏的方式进行结合指导DeAgent学习;把状态 s_{it} 输入到MILTM模型,将最后一层softmax函数之前输出的patch良性和恶性 $Z_{MILTM}^{s_{it}}$ 信息经过平滑之后作为软标签 $\hat{y}_{soft}^{s_{it}}$,将 s_{it} 对应的组织切片 X_i 的标签 Y_i 经过one-hot编码作为硬标签 $\hat{y}_{hard}^{s_{it}}$;其中,用Kullback-Leibler散度作为软损失衡量MILTM与DeAgent预测良性和恶性概率的差异,用交叉熵损失函数作为硬损失计算DeAgent的预测结果 $Z_{DeAgent}^{s_{it}}$ 与真实标签之间的误差;设置超参数 ε 平衡软标签和硬标签在总损失中的比例;具体公式为:

$$[0037] \quad L_{DeAgent} = \varepsilon * H^2 * KLdiv(\hat{y}_{soft}^{s_{it}}, \hat{y}_{DeAgent}^{s_{it}}) + (1 - \varepsilon) * CE(\hat{y}_{hard}^{s_{it}}, \hat{y}_{DeAgent}^{s_{it}}) \quad (13)$$

$$[0038] \quad \hat{y}_{soft}^{s_{it}} = \frac{\exp(\frac{Z_{MLTM}^{s_{it},k}}{H})}{\sum_{k=1}^K \exp(\frac{Z_{MLTM}^{s_{it},k}}{H})} \quad (14)$$

$$[0039] \quad \hat{y}_{DeAgent}^{s_{it}} = \frac{\exp(\frac{Z_{DeAgent}^{s_{it},k}}{H})}{\sum_{k=1}^K \exp(\frac{Z_{DeAgent}^{s_{it},k}}{H})} \quad (15)$$

[0040] 其中,K是类别,H是温度系数,用H来调节软标签 $\hat{y}_{soft}^{s_{it}}$ 中包含的信息。

[0041] 并且,基于强化学习的病理切片扫描分析一体化系统,包括数字病理自动扫描系统和人工智能辅助分析系统。以将扫描、传入计算机、人工智能辅助处理这一阅片顺序过程,转变为扫描与人工智能辅助处理同步进行,进而实现病理切片扫描分析一体化,减少阅片所用时间。

[0042] 所述数字病理自动扫描系统包括图像获取模块、硬件控制模块、图像处理模块和数据管理模块;

[0043] 所述图像获取模块由光源、显微镜、相机组成,用于采集切片清晰的局部显微图像;

[0044] 所述硬件控制模块,可根据切片扫描的需求,控制运动平台到达指定位置;

[0045] 所述图像处理模块用于完成局部显微图像的预处理,与上述硬件控制模块配合采集清晰的局部显微图像,将局部显微图像配准、融合成为全景数字切片;

[0046] 所述数据管理模块分为参数存储和图像数据管理两部分;其中,参数存储部分用于记录运动平台的运动参数和扫描参数,图像管理部分用于保存和载入显微图像。

[0047] 所述人工智能辅助分析系统,利用设计好的人工智能算法,对所述病理图像进行分析,包括图片采集传输模块、图像预处理模块、分析模块、可视化模块;

[0048] 所述图像采集传输模块用于将镜头拍摄的图片传输给人工智能系统进行分析;

[0049] 所述图像预处理模块用于将传入的图片切割成适合的小块,以便后续的基于人工智能的分析;

[0050] 所述分析模块通过模型对输入图片的类别、是否存在异常区域、异常区域的位置等进行分析并输出;

[0051] 所述可视化模块用于当整张病理图扫描并分析结束后输出原病理图及其上的异常区域的框取,并可将异常区域进行放大观察。

[0052] 本发明的有益效果:该基于强化学习的病理切片扫描分析一体化方法及系统包括用于获取、采集切片清晰的局部显微图像并对局部显微图像进行预处理的数字病理扫描系统,以将局部显微图像配准、融合成为全景数字切片,并将处理后的显微图像以及相关参数保存;还包括用于将数字病理扫描系统所获得的图像进行分析的人工智能辅助分析系统,进而判断是否存在异常区域,并且,将异常区域标记并可视化输出。该一体化方法及系统通过模型模仿病理专家的诊断过程,首先在低分辨率整张幻灯片图像中采样可疑的疾病区域,然后切换到高倍率下进一步确认;并通过卷积神经网络和循环神经网络对选定的感兴趣区域进行处理,学习不同疾病区域的判别特征,并且预测下一个位置。能够使组织玻片扫

描过程与人工智能分析并行,可简化人力物力,病理判断效率高,实用性强。并且,本发明能够在皮肤病理图像黑色素瘤和良性痣的分类任务中取得良好的检测精度。

附图说明

[0053] 图1是本发明的工作流程示意图。

[0054] 图2是本发明面向智能搜索的深度强化学习训练过程框图。

具体实施方式

[0055] 为了解决目前计算机辅助诊断在病理应用中存在的技术问题,本发明提出一种基于强化学习的病理切片扫描分析一体化方法及系统。病理专家在诊断时,通常不会在最高分辨率放大镜下分析幻灯片的每个部分,而是先在低分辨率下找到一些可疑区域,进而再切换到高分辨率下进一步确认,此过程可以看作是一个有顺序的马尔科夫决策。该一体化方法及系统通过模型模仿病理专家的诊断过程:首先在低分辨率整张幻灯片图像中采样可疑的疾病区域,然后切换到高倍率下进一步确认;并通过卷积神经网络和循环神经网络对选定的感兴趣区域进行处理,学习不同疾病区域的判别特征,并且预测下一个位置。模型无需对所有给定的样本进行学习,仅仅对一些高度可疑的疾病区域进行学习,在得到较高准确率的同时,可提高模型的速度,对于病理计算机辅助具有重要意义。

[0056] 详细说明本发明方法的具体步骤。该基于强化学习的病理切片扫描分析一体化方法包括:

[0057] 步骤一、从医院的皮肤病理数据库中,通过病理数字切片扫描仪收集数据,病理专家对数据进行进一步的确认,以保证数据标签的准确性。

[0058] 步骤二、将收集的病理数据集,经过数据预处理,输入到基于卷积神经网络搭建的多实例专家指导模型,并采用交叉熵损失函数对专家指导模型进行训练。进而,经过不断的训练来获得一个能够指导深度强化学习的智能体进行决策的专家指导模型。

[0059] 所建立的多实例专家指导模型如公式(1)所示:

$$[0060] \quad y = f_{MILTM_{Res}}^{\theta}(x) \quad (1)$$

[0061] 其中,x是经过专家视点约束的病理图像数据, $f_{MILTM_{Res}}^{\theta}$ 是具有可训练参数 θ 的多实例专家指导模型,使用ResNet卷积网络实现,y是卷积神经网络输出的预测结果;

[0062] 用于指导模型训练的交叉熵损失函数如公式(2)所示:

$$[0063] \quad L = \frac{1}{N} \sum_{i=1}^N L_i = \frac{1}{N} \sum_{i=1}^N -[Y_i \cdot \log(y_i) + (1 - Y_i) \cdot \log(1 - y_i)] \quad (2)$$

[0064] 其中, Y_i 是第i张切片的标签, y_i 是模型输出的第i张切片的预测结果。

[0065] 步骤三、将收集的病理数据集,分为训练集、验证集和测试集。当训练深度强化学习模型时,对训练集中的图像依次进行处理,设置迭代次数,限制每一张图像被迭代的次数。利用贪婪策略和动作网络获取智能体的动作,进而生成训练样本,并且,输入到基于多实例学习构建的专家模型和强化学习模型中。

[0066] 动作网络被定义为搜索智能体,用于输出动作,具体来说:使用长短记忆网络(long shortterm memory,LSTM)与全连接层结合设计了搜索智能体—SeAgent;并通过优

化参数 $\varphi_{SeAgent}$ 使SeAgent快速查找可疑区域。

[0067] SeAgent输入的是状态 s_{it} 的特征向量 $v_{MILTM_Res}^{s_{it}}$ ； $v_{MILTM_Res}^{s_{it}}$ 表示 s_{it} 进入到MILTM经过卷积层、RULE非线性激活层、池化层、ResNet-34的四个block块(F_{MILTM}^{Res})输出的特征。使用预训练好的MILTM提取状态特征向量,不仅缩短了SeAgent特征提取的时间,而且提高了SeAgent的精准性。强化学习的最终目标是为了获得最大累积回报,SeAgent的进化需要依赖于过去和前状态的信息和奖励。相比于卷积神经网络和标准的循环神经网络,LSTM非常适合于长期时间序列数据的分类、处理和预测,并且能够很好的解决因网络深度的增加而造成梯度消失现象。

[0068] 利用公式(5)把 $v_{MILTM_Res}^{s_{it}}$ 经过LSTM($F_{SeAgent}^{LSTM}$)与FC($F_{SeAgent}^{FC1}$)的特征 $v_{SeAgent_LSTM}^{s_{it}}$ 、 $v_{SeAgent_FC1}^{s_{it}}$ 进行融合,使得SeAgent充分的考虑了当前状态特征与先前状态信息。将融合后的特征经过全连接层 $F_{SeAgent}^{FC2}$ 和softmax函数,最终输出SeAgent的下一步动作 $\pi_{\Phi_{SeAgent}}(s_{it}, a_{it})$ 公式(6)。

$$[0069] \quad v_{SeAgent_LSTM}^{s_{it}} = F_{SeAgent}^{LSTM}(v_{MILTM_Res}^{s_{it}} | \varphi_{SeAgent_LSTM}) \quad (3)$$

$$[0070] \quad v_{SeAgent_FC1}^{s_{it}} = F_{SeAgent}^{FC1}(v_{MILTM_Res}^{s_{it}} | \varphi_{SeAgent_FC1}) \quad (4)$$

$$[0071] \quad v_{fusion}^{s_{it}} = v_{SeAgent_LSTM}^{s_{it}} + v_{SeAgent_FC1}^{s_{it}} \quad (5)$$

$$[0072] \quad \begin{aligned} \pi_{\Phi_{SeAgent}}(a_{it} | s_{it}) &= P_{\Phi_{SeAgent}}(a_{it} | s_{it}, \varphi_{SeAgent}) \\ &= \text{softmax}(F_{SeAgent}^{FC2}(v_{fusion}^{s_{it}} | \varphi_{SeAgent_FC2})) \end{aligned} \quad (6)$$

[0073] 由于SeAgent不可微,使用策略梯度进行优化参数 $\varphi_{SeAgent}$,以学习最优搜索策略 $\pi_{\Phi_{SeAgent}}(a_{it} | s_{it})$ 实现 R_t 的期望最大化,定义目标函数为:

$$[0074] \quad L_{SeAgent} = E \left[\sum_{t=1}^T \gamma^{t-1} r_{it} \right] = \sum_{t=1}^T P_{\Phi_{SeAgent}}(a_{it} | s_{it}, \varphi_{SeAgent}) \gamma^{t-1} r_{it} \quad (7)$$

[0075] SeAgent每完成一次任务,对搜索网络进行一次参数更新;目标函数 $L_{SeAgent}$ 求梯度的公式为:

$$[0076] \quad \nabla_{\Phi_{SeAgent}} L_{SeAgent} = \sum_{t=1}^T \gamma^{t-1} r_{it} \nabla_{\Phi_{SeAgent}} \log(\pi_{\Phi_{SeAgent}}(a_{it} | s_{it})) \quad (8)$$

[0077] 步骤四、专家模型对输入的样本进行预测,其预测结果指导对强化学习智能体进行训练;采用策略梯度和奖励函数对智能体输出动作的网络进行训练,通过知识蒸馏的方式对智能体的决策分类网络进行训练;训练后的智能体生成下一个动作,作为新的样本输入到专家模型和强化学习模型,直到满足预先设置的终止条件,该图像的处理过程结束。

[0078] 为了训练智能体输出动作的SeAgent网络,根据任务特点设计奖励函数(reward);具体来说:奖励函数是驱动SeAgent选择最大奖励值动作的,使用公式(9),鼓励SeAgent高效快速的完成疾病诊断任务。如果SeAgent选择的动作获取的状态 s_{it} 经过MILTM鉴别为肿瘤,则反馈给SeAgent的+2的奖励,否则给予一个-2的奖励。此外,如果 s_{it}' 也被MILTM鉴别为肿瘤,则将SeAgent的奖励扩大到5;奖励是在每个行动之后立即产生的。

$$[0079] \quad r_{it} = \begin{cases} -2 & \text{if } \hat{y}_{MILTM}^1 = 0 \\ +2 & \text{if } \hat{y}_{MILTM}^1 = 1 \text{ and } \hat{y}_{MILTM}^2 = 0 \\ +5 & \text{if } \hat{y}_{MILTM}^1 = 1 \text{ and } \hat{y}_{MILTM}^2 = 1 \end{cases} \quad (9)$$

[0080] \hat{y}_{MILTM}^1 和 \hat{y}_{MILTM}^2 表示在状态 s_{it} 和 s_{it}' 时, F_{MILTM}^1 和 F_{MILTM}^2 预测的结果;其中:

$$[0081] \quad \hat{y}_{MILTM}^1 = \operatorname{argmax}(F_{MILTM}^1(s_{it} | \theta_{MILTM}^1)) \quad (10)$$

$$[0082] \quad \hat{y}_{MILTM}^2 = \operatorname{argmax}(F_{MILTM}^2(s_{it}' | \theta_{MILTM}^2)) \quad (11)$$

[0083] 最终,通过公式 (12), 计算获得最大的累积回报来指导SeAgent进行优化, γ^{t-1} 是 r_{it} 在 t 时刻奖励的权重因子。

$$[0084] \quad R_i = \sum_{t=1}^T \gamma^{t-1} r_{it} \quad (12)$$

[0085] 同时,对于决策智能体DeAgent部分,使用一层全连接构建;具体来说:通过优化全连接层 $F_{DeAgent}$ 的参数 $\psi_{DeAgent}$, 实现对状态 s_{it} 特征向量 $v_{MILTM_Res}^{s_{it}}$ 的精准分类。由于病理组织图像种包含大量的非肿瘤组织区域,在没有像素级注释的情况下,仅使用组织图像的全局标签和交叉熵损失函数很难实现对DeAgent的优化。此外,SeAgent每一步的状态 s_{it} , 获取的 patch 的标签不一定与全局标签保持一致。

[0086] 并且,受教师-学生模型的启发,将交叉熵损失函数与知识蒸馏的方式进行结合指导DeAgent学习。把状态 s_{it} 输入到MILTM模型,将最后一层softmax函数之前输出的patch良性和恶性 $Z_{MILTM}^{s_{it}}$ 信息经过平滑之后作为软标签 $\hat{y}_{soft}^{s_{it}}$, 将 s_{it} 对应的组织切片 X_i 的标签 Y_i 经过 one-hot 编码作为硬标签 $\hat{y}_{hard}^{s_{it}}$ 。

[0087] 其中,用Kullback-Leibler (KL) 散度作为软损失衡量MILTM与DeAgent预测良性和恶性概率的差异,用交叉熵损失函数作为硬损失计算DeAgent的预测结果 $Z_{DeAgent}^{s_{it}}$ 与真实标签之间的误差。设置超参数 ε 平衡软标签和硬标签在总损失中的比例; ε 加权系数越大,表明诱导DeAgent依赖于MILEG的贡献,这对于训练初期阶段指导DeAgent识别良性和恶性组织是至关重要的。

[0088] 具体公式为:

$$[0089] \quad L_{DeAgent} = \varepsilon * H^2 * KLdiv(\hat{y}_{soft}^{s_{it}}, \hat{y}_{DeAgent}^{s_{it}}) + (1 - \varepsilon) * CE(\hat{y}_{hard}^{s_{it}}, \hat{y}_{DeAgent}^{s_{it}}) \quad (13)$$

$$[0090] \quad \hat{y}_{soft}^{s_{it}} = \frac{\exp(\frac{Z_{MILTM}^{s_{it}_k}}{H})}{\sum_{k=1}^K \exp(\frac{Z_{MILTM}^{s_{it}_k}}{H})} \quad (14)$$

$$[0091] \quad \hat{y}_{DeAgent}^{s_{it}} = \frac{\exp(\frac{Z_{DeAgent}^{s_{it}_k}}{H})}{\sum_{k=1}^K \exp(\frac{Z_{DeAgent}^{s_{it}_k}}{H})} \quad (15)$$

[0092] 其中, K 是类别, H 是温度系数, 用 H 来调节软标签 $\hat{y}_{soft}^{s_{it}}$ 中包含的信息。经过 H 调节的软标签, 可以把原本模型输出的信息在维持差异的同时保留低概率样本的信息, 减少分类错

误的概率。

[0093] 步骤五、返回步骤二,继续从训练集中输入下一个图像,直到训练集中的所有图像处理结束,强化学习网络参数确定。

[0094] 用于病理图像快速判断的深度强化学习网络,包括弱监督深度强化学习模型,该模型可以对病理图像缺乏进行标注,从而避免高成本的手动标注过程。

[0095] 步骤六、通过训练好的深度强化学习模型,对测试集中的图像进行疾病检测和分类,输出分类结果。

[0096] 另外,基于强化学习的病理切片扫描分析一体化系统,包括数字病理自动扫描系统和人工智能辅助分析系统;进而将扫描、传入计算机、人工智能辅助处理这一阅片顺序过程,转变为扫描与人工智能辅助处理同步进行,进而实现病理切片扫描分析一体化,减少阅片所用时间。

[0097] 数字病理自动扫描系统包括图像获取模块、硬件控制模块、图像处理模块和数据管理模块。其中,图像获取模块由光源、显微镜、相机组成,用于采集切片清晰的局部显微图像。硬件控制模块可根据切片扫描的需求,控制运动平台到达指定位置。图像处理模块用于完成局部显微图像的预处理,与上述硬件控制模块配合采集清晰的局部显微图像,将局部显微图像配准、融合成为全景数字切片。数据管理模块分为参数存储和图像数据管理两部分,其中,参数存储部分用于记录运动平台的运动参数和扫描参数,图像管理部分用于保存和载入显微图像。

[0098] 人工智能辅助分析系统利用设计好的人工智能算法,对病理图像进行分析,包括图片采集传输模块、图像预处理模块、分析模块、可视化模块。图像采集传输模块用于将镜头拍摄的图片传输给人工智能系统进行分析。图像预处理模块用于将传入的图片切割成适合的小块,以便后续的基于人工智能的分析。分析模块通过模型对输入图片的类别(表皮,真皮,皮下)、是否存在异常区域、异常区域的位置等进行分析并输出。可视化模块用于当整张病理图扫描并分析结束后输出原病理图及其上的异常区域的框取,并可将异常区域进行放大观察。

[0099] 该基于强化学习的病理切片扫描分析一体化系统的数字病理自动扫描系统,用于获取、采集切片清晰的局部显微图像并对局部显微图像进行预处理,将局部显微图像配准、融合成为全景数字切片,最后将上述处理后的显微图像以及相关参数保存。人工智能辅助分析系统,则用于将数字病理扫描系统所获得的图像进行分析,判断是否存在异常区域,并且,将异常区域标记并可视化输出。

[0100] 如表1所示,该基于强化学习的病理切片扫描分析一体化方法中的深度强化学习网络与MIL-VGG、MIL-ResNet对比可得,本方法中深度强化学习网络的准确性略强于现有主流网络。

[0101] 表1本方法中深度强化学习网络的各项准确率与常用网络的对比表

[0102]

Methods		Precision	Specificity	F1	ACC
10x	MIL-VGG19	0.833±0.178	0.800±0.118	0.897±0.063	0.887±0.095
	MIL-VGG16	0.827±0.060	0.787±0.085	0.904±0.035	0.893±0.043
	MIL-ResNet50	0.853±0.047	0.825±0.065	0.920±0.027	0.912±0.032
	MIL-ResNet34	0.863±0.071	0.837±0.085	0.919±0.026	0.912±0.032
5x	MIL-VGG19	0.873±0.138	0.833±0.202	0.928±0.082	0.916±0.101
	MIL-VGG16	0.858±0.087	0.825±0.126	0.921±0.052	0.907±0.057
	MIL-ResNet50	0.815±0.077	0.775±0.106	0.900±0.043	0.887±0.053
	MIL-ResNet34	0.885±0.061	0.867±0.076	0.938±0.034	0.930±0.038
Ours		0.950±0.034	0.953±0.031	0.927±0.009	0.932±0.005

[0103] 如表2所示,该基于强化学习的病理切片扫描分析一体化方法中的深度强化学习网络与MIL-VGG、MIL-ResNet对比可得,本方法中深度强化学习网络的训练时长远低于现有主流网络。

[0104] 表2本方法中深度强化学习网络的训练时长与常用网络的对比表

[0105]

Methods		Run Time (s)
10x	MIL-VGG19	4568.14±23.388
	MIL-VGG16	4126.89±12.637
	MIL-ResNet50	3264.01±11.521
	MIL-ResNet34	3282.13±2.916
5x	MIL-VGG19	4250.66±3.449
	MIL-VGG16	3843.99±2.754
	MIL-ResNet50	2814.29±2.355
	MIL-ResNet34	2831.50±2.613
Ours		723.68±11.970

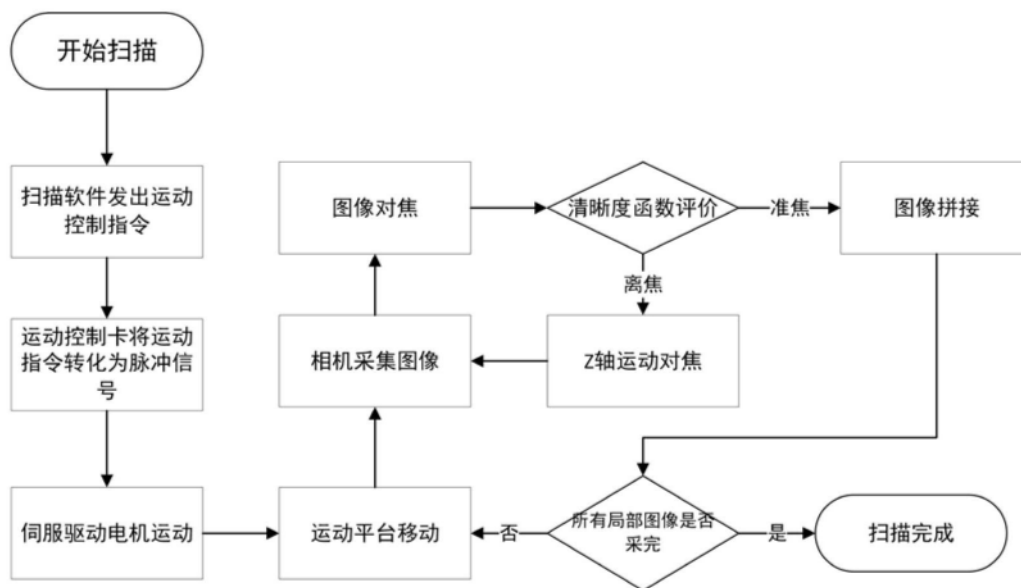


图1

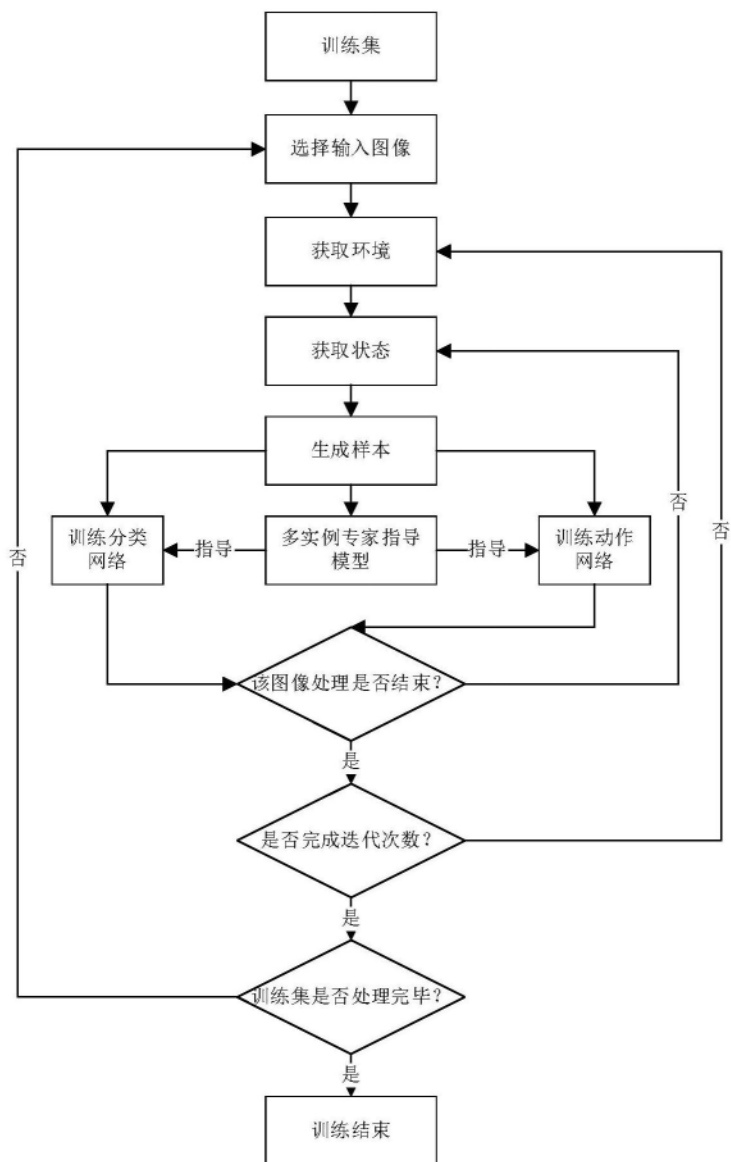


图2