



(12)发明专利申请

(10)申请公布号 CN 110946552 A

(43)申请公布日 2020.04.03

(21)申请号 201911041423.7

(22)申请日 2019.10.30

(71)申请人 南京航空航天大学

地址 210017 江苏省南京市秦淮区御道街
29号

(72)发明人 钱志余 刘文文 晋晓飞 李俊俊
张欢

(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 曹芸

(51)Int.Cl.

A61B 5/00(2006.01)

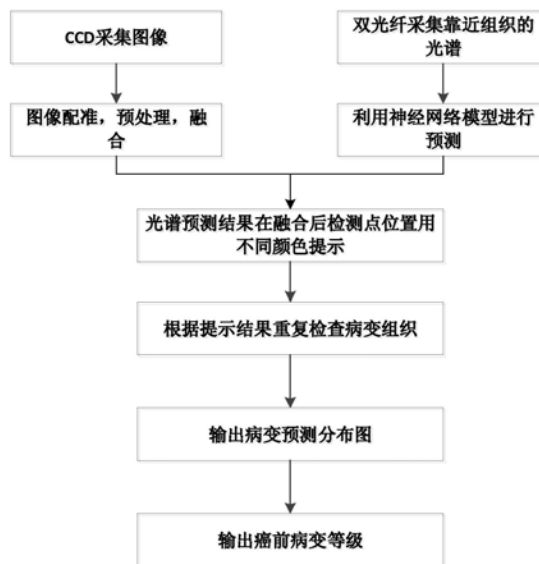
权利要求书1页 说明书5页 附图4页

(54)发明名称

一种结合光谱和图像的宫颈癌前病变筛查方法

(57)摘要

本发明涉及一种结合光谱和图像的宫颈癌前病变筛查方法,属于光谱分析和医学图像处理领域。本发明根据光谱和内窥镜的联合探头同时采集到的光谱和图像数据构建数据库,利用光谱数据库构建神经网络癌前病变分类模型,利用图像配准和融合,将对应检测点光谱诊断结果用不同颜色显示在对应检测点图像,重构成癌前病变分布图,根据分布图输出诊断结果。本发明可以单独作为宫颈癌筛查算法,也可以为进一步活检提供图像向导。本方法的诊断结果为一张宫颈癌前病变分布图,客观直接,不受操作人员的技术影响。



1. 一种结合光谱和图像的宫颈癌前病变筛查方法,其特征在于:包括以下几个步骤:

(1) 数据采集:使用数据采集系统,进行宫颈组织光谱和图像的在体采集;

(2) 数据预处理:宫颈组织的原始光谱和图像信息预处理;

(3) 建立数据库:预处理后的数据进行特征提取,并保存到数据库中;

(4) 模型训练:模型训练使用BP神经网络,提取光谱特征作为输入,以0和1作为输出,分别属于训练集的两类;

(5) 结果预测:采用一对一的训练方式,共十个分类器,最终结果由十个分类器投票决定,得票最多的类别即为预测结果。

2. 根据权利要求1所述的一种结合光谱和图像的宫颈癌前病变筛查方法,其特征在于:步骤(1)中所述数据采集系统包括光源、光谱仪、联合探头和CCD相机,其中联合探头分别与光源、CCD相机和光谱仪连接。

3. 根据权利要求2所述的一种结合光谱和图像的宫颈癌前病变筛查方法,其特征在于:所述联合探头包括光谱光纤、导光光纤和成像光纤,三种光纤独立包裹,整合在同一根探头中。

4. 根据权利要求1所述的一种结合光谱和图像的宫颈癌前病变筛查方法,其特征在于:步骤(1)的具体过程如下:

使用数据采集系统的探头进行光谱和图像的采集,首先在距离移行带较远的位置拍摄子宫颈的整体图像作为定位基准,后紧贴子宫颈组织进行局部图像采集以及光谱采集;疑似病变区域重点采集,探头绕移行带一周,采集完毕。

5. 根据权利要求1所述的一种结合光谱和图像的宫颈癌前病变筛查方法,其特征在于:步骤(2)的具体过程如下:

a) 原始光谱数据,每10个点取平均进行平滑处理,然后做归一化;

b) 采集的图像进行压缩和特征提取,然后按照时间序列进行配准和融合。

一种结合光谱和图像的宫颈癌前病变筛查方法

技术领域

[0001] 本发明涉及一种结合光谱和图像的宫颈癌前病变筛查方法,属于光谱分析和医学图像处理领域。

背景技术

[0002] 宫颈癌是一种常见的妇科恶性肿瘤,对女性的健康造成了严重的危害。宫颈癌早期的治愈率可达90%以上,而原位癌的治愈率为100%。宫颈癌的早期筛查可以防止宫颈癌的继续发展,预防宫颈癌的发生。因此,宫颈癌的早期诊断至关重要。临床上根据子宫颈上皮细胞病变程度将宫颈癌早期病变分为CINI级、CIN II级和CIN III级,CIN III级以后就是浸润癌。早期病变到浸润癌的发展缓慢,及时和普遍的早期筛查是阻断宫颈癌的有效手段。

[0003] 近年来临床上宫颈癌的早期筛查有各种方法,如光学、电学和各种诊断试剂,也可以通过疫苗来预防宫颈癌。临床上现有的检测方法分为观察法、细胞学检测、组织病理学检查等几类。

[0004] 临床上使用最广泛的一种方法是巴氏涂片,巴氏涂片指宫颈脱落细胞涂片,是指从子宫颈部取少量的细胞样品,放在玻璃片上,然后在显微镜下研究是否异常。这种方法造价低,在基层也可以应用,但是特异度不高。目前临床上还缺少一种灵敏度和特异性良好且价格友好的宫颈癌筛查方法,以应用到医疗资源不发达的基层医院,达到提高宫颈癌筛查率和检查准确率的目的。

[0005] 宫颈癌在发展过程中,早期病变分为三个阶段,CINI级、CIN II级和CIN III级。当上皮层的下 $\frac{1}{3}$ 区被异型细胞占满时称为CIN I。当上皮层的 $\frac{1}{3}$ — $\frac{2}{3}$ 被异型细胞占满时称为CIN II,异型性较CIN I级明显。异型细胞扩散面积超过上皮层的 $\frac{2}{3}$ 区时称为CIN III级。

[0006] 研究发现宫颈发育异常中,上皮细胞核质比和约化散射系数增大,基质中新生血管增多,胶原网络的降解,吸收系数增大,约化散射系数减小,散射减少可归因于宫颈基质中胶原网络的破坏。吸收系数和约化散射系数的变化可以利用组织的光谱进行计算。

[0007] 神经网络算法对模糊的、非线性的、甚至是含有噪声的样本进行分类或者预测。反向传播(Back Propagation, BP)神经网络采用误差逐层向后传播的方式调节每层参数权重,最终获得符合要求的网络误差平方和。

发明内容

[0008] 本发明提出了一种结合光谱和图像的宫颈癌前病变筛查方法,能够将光谱和图像信息定位到某检测点,根据建立的网络模型判断癌前病变情况,利用图像融合技术,用不同颜色将病变信息反映在子宫颈图上,并输出诊断结果。

[0009] 本发明为解决其技术问题采用如下技术方案:

[0010] 一种结合光谱和图像的宫颈癌前病变筛查方法,包括以下几个步骤:

[0011] (1) 数据采集:使用数据采集系统,进行宫颈组织光谱和图像的在体采集;

- [0012] (2) 数据预处理: 宫颈组织的原始光谱和图像信息预处理;
- [0013] (3) 建立数据库: 预处理后的数据进行特征提取, 并保存到数据库中;
- [0014] (4) 模型训练: 模型训练使用BP神经网络, 提取光谱特征作为输入, 以0和1作为输出, 分别属于训练集的两类;
- [0015] (5) 结果预测: 采用一对一的训练方式, 共十个分类器, 最终结果由十个分类器投票决定, 得票最多的类别即为预测结果。
- [0016] 步骤(1)中所述数据采集系统包括光源、光谱仪、联合探头和CCD相机, 其中联合探头分别与光源、CCD相机和光谱仪连接。
- [0017] 所述联合探头包括光谱光纤、导光光纤和成像光纤, 三种光纤独立包裹, 整合在同一根探头中。
- [0018] 步骤(1)的具体过程如下:
- [0019] 使用数据采集系统的探头进行光谱和图像的采集, 首先在距离移行带较远的位置拍摄子宫颈的整体图像作为定位基准, 后紧贴子宫颈组织进行局部图像采集以及光谱采集; 疑似病变区域重点采集, 探头绕移行带一周, 采集完毕。
- [0020] 步骤(2)的具体过程如下:
- [0021] a) 原始光谱数据, 每10个点取平均进行平滑处理, 然后做归一化。
- [0022] b) 采集的图像进行压缩和特征提取, 然后按照时间序列进行配准和融合。
- [0023] 本发明的有益效果如下:
- [0024] 本发明提供了一种宫颈癌前病变的筛查方法, 基于光谱的特征参数, 用BP神经网络构建分类模型, 输出分类结果。本方法采用了科学的数据采集方式, 能够得到全面的宫颈病变分布状况, 无需对宫颈进行刺激, 安全可靠。经大量实验数据验证, 本方法的判别准确率在85%以上, 在已有的方法中具有相当的优越性, 在宫颈癌筛查领域是一项重大突破。

附图说明

- [0025] 图1是本发明整体算法的流程示意图。
- [0026] 图2是本发明神经网络建模的流程示意图。
- [0027] 图3是本发明神经网络参数优化的流程示意图。
- [0028] 图4是本发明图像融合的流程示意图。

具体实施方式

- [0029] 下面结合附图对本发明创造做进一步详细说明。
- [0030] 本发明涉及一种宫颈癌前病变筛查的方法, 利用光谱提取的特征建立神经网络模型进行癌前病变的分级, 利用图像融合技术进行子宫颈口移行带的定位。将光谱检测结果显示在移行带图像上, 不同级别用不同颜色等级表示, 给医务工作者更加直观的展示, 并为病变区域重点检查和活检提供向导。
- [0031] 本方法中所使用的数据共分为五类, 其标签分别是正常、CINI、CIN II、CIN III和宫颈糜烂。数据来源于医院临床试验, 标签取自组织病理学检验结果。
- [0032] 图1为采用的筛查算法的整体流程示意图。包括:
- [0033] CCD(电荷耦合器件)图像和光谱仪同步采集的数据, 数据采集探头在距离宫颈口

移行带较远处先采集移行带的整体图像,之后慢慢靠近移行带,在接触到移行带后停止前进。此过程中,CCD图像被筛选并保存用作图像融合的基准;光谱数据根据筛查算法,去掉不符合检测距离的光谱数据,只保留有效数据。之后探头沿移行带表面螺旋向内滑动,采集图像和光谱数据。

[0034] 原始光谱数据和预处理后的数据都进行参数提取,分别计算波长在200-1100nm之间几个特征波段的光谱斜率,和光谱峰下面积,以及吸收系数和约化散射系数两个光学参数。疑似病变部位的组织病理学检查结果作为本算法数据的标签来源。以上数据均保存在数据库中。

[0035] 根据采集的图像进行压缩和特征提取,然后按照时间序列进行配准和融合,示意图见4,并在移行带整体图像上定位。

[0036] 根据已经建好的神经网络模型进行检测,根据时间序列记录检测结果。模型的建立见图2所示示意图,模型的优化选择见图3所示。

[0037] 根据时间序列的图像定位,根据光谱检测结果的癌前病变分级将结果用不同颜色表示,在融合后的图像定位区域显示不同颜色。形成一张癌前病变等级的实时分布图。

[0038] 根据定位区域的等级提示,重点重复检查不同CIN等级的区域及周边区域,以此提高检测准确性和重复性,并且确定病变范围。

[0039] 根据区域提示,对未检查的区域补充检查,输出最终区域判断,形成一张完整的癌前病变分布图。

[0040] 根据病变分布,确定不同病变等级的分布面积,根据病变面积的大小进行等级判断。此过程可以适用阈值判断,超过阈值即判断为对应的等级;也可以再次建立神经网络模型,输出判断结果。

[0041] 图2是本发明的神经网络的流程示意图。

[0042] 一种用于判断癌前病变的等级的神经网络模型建模方法,建立宫颈癌早期光谱数据库,将数据集分为训练集和测试集,利用训练集调节网络参数,通过测试集选择最优模型,用于诊断分析。本方法采用的神经网络模型,以S形转移函数作为激活函数,11维特征为输入,以0和1的方式输出分类结果。

[0043] 采用单隐藏层神经网络,隐藏层节点数的确定利用经验公式 $m = \sqrt{n+l} + \alpha$, 其中 α 为1-10之间的常数,m为隐藏层节点数,n为输出层输入层节点数,l为输出层节点数。设置隐藏层神经单元为3-13个,误差分别为0.1和0.01分别进行建模试验,最终得出隐藏层神经单元数为10,误差为0.01时,所建模型准确率最高。

[0044] 随机抽取正常组织和肿瘤组织作为训练集,剩余正常组织和肿瘤组织样本作为验证集进行神经网络模型训练和验证。将 μ'_s 、 μ_a 、S₅₀₀₋₅₂₀、S₅₂₄₋₅₃₂、S₅₄₀₋₅₆₀、S₅₆₅₋₅₇₀、S₅₇₅₋₅₉₀、S₇₅₀₋₈₅₀、A₄₀₀₋₈₀₀、A₅₀₀₋₅₅₀、A₅₅₀₋₆₀₀等11维特征作为输入,将肿瘤组织作为正样本,正常组织作为负样本,正常组织和肿瘤组织分别标记为0和1,作为输出。

[0045] 本方法使用的光谱特征参数,在裸鼠试验中的其中一组样例如表1所示:

[0046] 表1裸鼠的正常和肿瘤组织光谱参数

[0047]	正常组织	肿瘤组织
μ'_s	11.22328	16.45858

μ_a	0.138665	0.1496
S ₅₀₀₋₅₂₀	385.5759	343.4692
S ₅₂₄₋₅₃₂	152.7547	-314.872
S ₅₄₀₋₅₆₀	572.737	405.058
S ₅₆₅₋₅₇₀	126.5193	-337.491
S ₅₇₅₋₅₉₀	278.7903	623.9977
A ₅₀₀₋₅₅₀	2.851088	1.882809
A ₅₅₀₋₆₀₀	6.832492	4.143716
S ₇₅₀₋₈₅₀	-171.7	-334.8
A ₄₀₀₋₈₀₀	212.027	131.512

[0048] 光谱在540nm和575nm处有明显的吸收峰,吸收峰处的斜率和面积可以作为区分不同病变组织的特征因子。选择明显接近直线段分布波段,求取斜率和面积。分别计算正常病人、CIN I期病人、CIN II期病人、CIN III期病人和宫颈糜烂病人的500-520nm、524-532nm、540-560nm、565-570nm和575-590nm波长范围的光谱斜率,随着癌前病变程度加深,各段斜率先减小后增大,说明斜率信息具有一定的识别意义,可以用作光谱特征。面积公式见(1),斜率面积见(2)。

$$[0049] \quad A = \sum_{i=1}^{n-1} \frac{y_{i+1} + y_i}{2} (x_{i+1} - x_i) \quad (1)$$

$$[0050] \quad S = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad (2)$$

[0051] 其中:A为面积, x_i 为波长, x_{i+1} 为下一个采集波长, y_i 为 x_i 这一波长对应的光强, y_{i+1} 为 x_{i+1} 这一波长对应的光强, n 为所求波段内的点的总数, S 为斜率;

[0052] 采用一对一策略,可以构建10个分类器。首先从中任选两个类作为正例和反例,用训练集进行建模,查看单个模型对训练集的分类效果,对于分类效果好的模型保留,效果不好的模型重复建模,直至效果较好。10个分类模型全都建好后,分别对所有的训练集和验证集样例按每个模型进行分类,记录每个样例属于哪一类,按照得票积分的形式合并10个分类器的预测结果,最终积分高的类作为最终的预测类。

[0053] 具体步骤如下:

[0054] 根据初始化的网络结构(包括层数、各层节点数、精度、学习率、步长、迭代次数等),利用归一化的训练集进行训练,计算隐藏层输出,计算输出层输出,根据每个位点训练结果和对应位点医院TCT(薄层液基细胞学)检查的标准结果进行对比,计算输出层偏差。

[0055] 判断是否满足循环条件,不满足时根据偏差调整公式分别调整输出层权值和隐藏层权值,再次进行循环计算。直到满足循环条件,保存各层权值到文档。

[0056] 利用测试集和已经读取的各层权值计算模型输出,根据灵敏度和特异性测试模型优劣。

- [0057] 新病患数据直接利用读取的各层权值进行输出计算,结果作为预测结果。
- [0058] 对于最优的神经网络模型选择,采用图3所示的神经网络参数优化方法,提高算法灵敏度和特异性。
- [0059] 设定网络模型限制条件,包括最大层数、最多节点数。
- [0060] 初始化网络参数,包括层数、节点数、精度、步长、学习率、循环次数。
- [0061] 将光谱数据库的数据按照不同比例分为训练集和测试集,分别用于神经网络的训练和测试。
- [0062] 利用训练集训练神经网络并保存权值等参数,利用剩余比例的测试集进行测试,分别保存测试结果。
- [0063] 迭代增加层数和节点数,继续进行模型训练和测试,直到满足限制条件。根据所有测试结果,选择特异性和灵敏度最高的模型。
- [0064] 图4是本发明的图像融合的流程示意图。
- [0065] 一种图像融合进行定位的方法,以整体宫颈图片为基准,将探头采集的不同检测点的图片进行配准和预处理后进行融合,然后根据不同检测点的诊断结果在融合后的图像上用不同等级表示。
- [0066] 在探头靠近子宫颈口的过程中,CCD由远及近拍摄移行带图像,并把移行带整体图作为基准。
- [0067] 探头在移行带表面移动时,按时间序列拍摄图片,对图片进行压缩,和特征提取。
- [0068] 根据特征进行图像配准,利用配准的图像进行融合,形成一张完整的探头运动轨迹图,以此来提示医生检测点位置,根据光谱的预测结果,可以对病变严重区域及周边进行重复检查,形成直观的病变分布图。

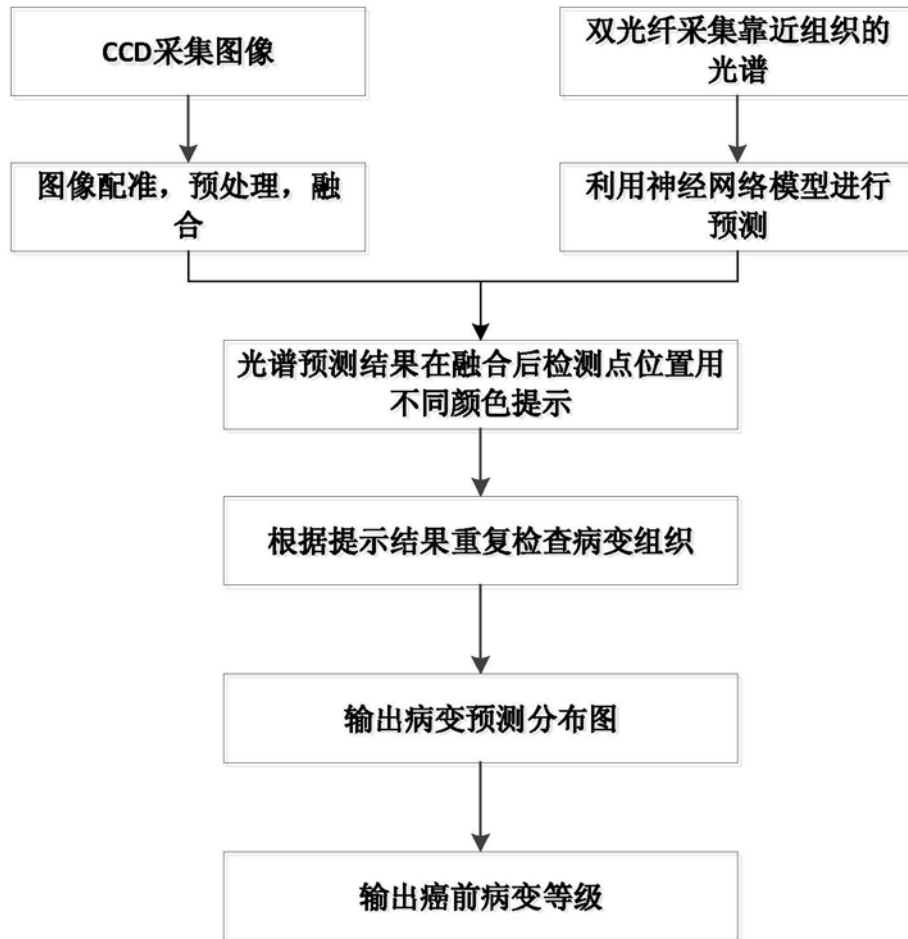


图1

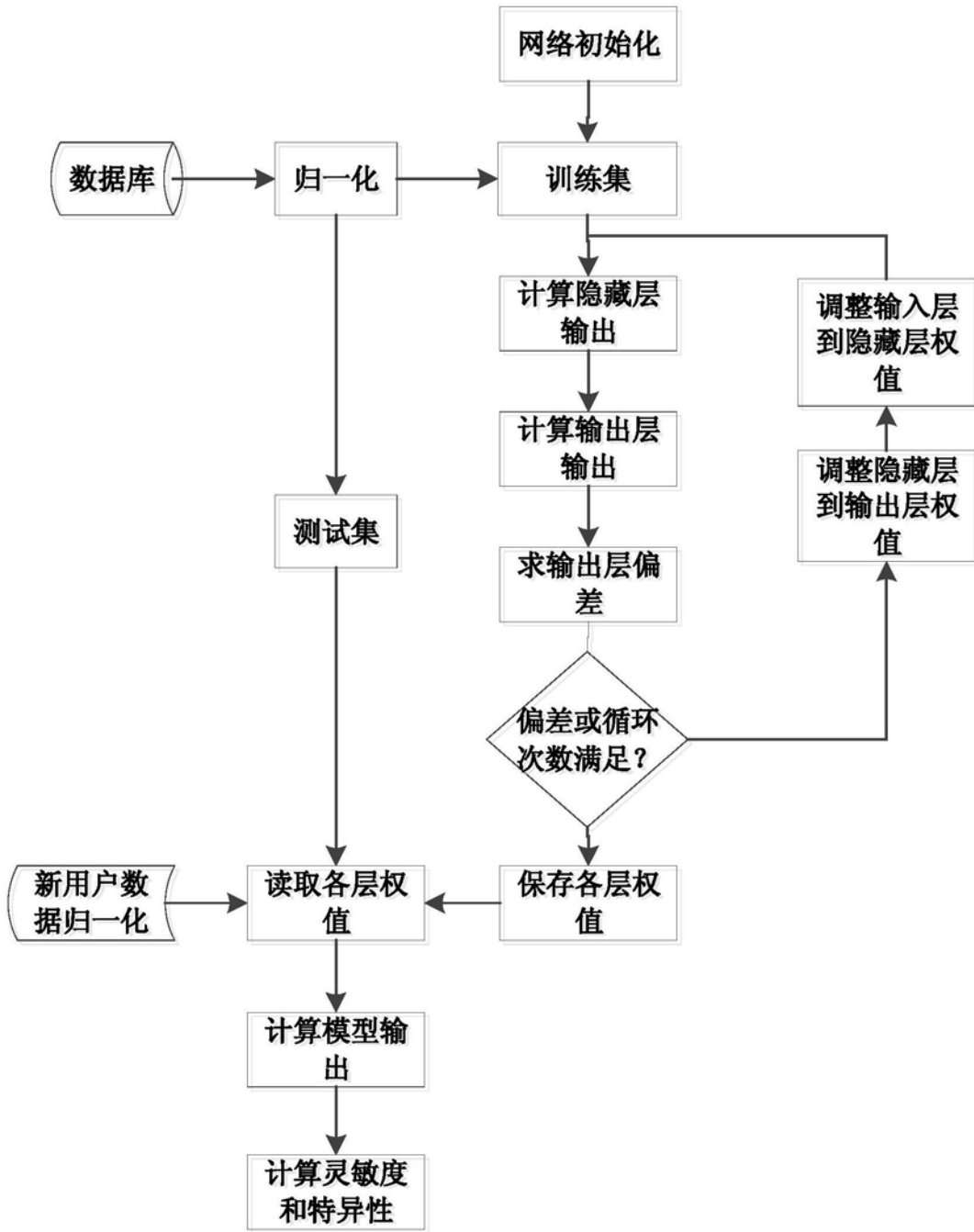


图2

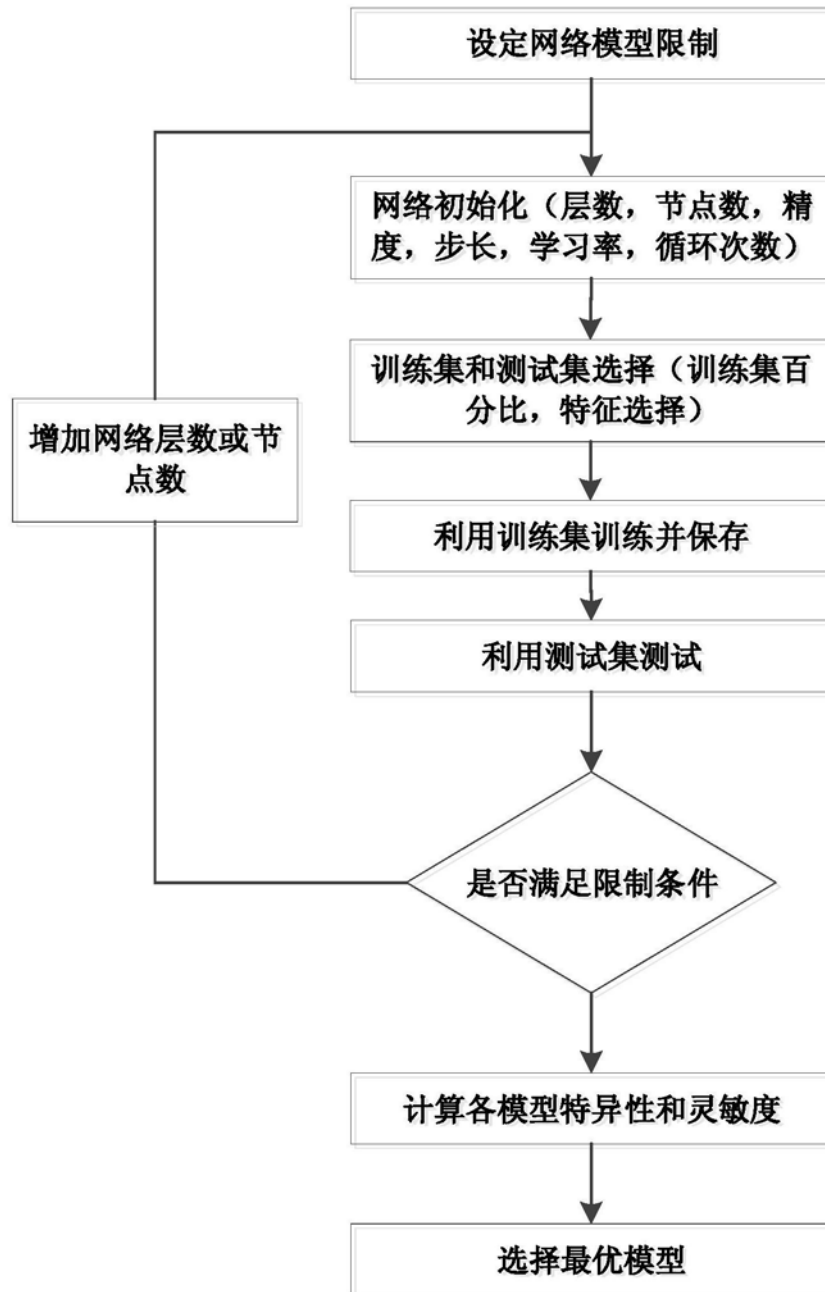


图3



图4

专利名称(译)	一种结合光谱和图像的宫颈癌前病变筛查方法		
公开(公告)号	CN110946552A	公开(公告)日	2020-04-03
申请号	CN201911041423.7	申请日	2019-10-30
[标]申请(专利权)人(译)	南京航空航天大学		
申请(专利权)人(译)	南京航空航天大学		
当前申请(专利权)人(译)	南京航空航天大学		
[标]发明人	钱志余 刘文文 晋晓飞 李俊俊 张欢		
发明人	钱志余 刘文文 晋晓飞 李俊俊 张欢		
IPC分类号	A61B5/00		
CPC分类号	A61B5/7264		
代理人(译)	曹芸		
外部链接	Espacenet SIPO		

摘要(译)

本发明涉及一种结合光谱和图像的宫颈癌前病变筛查方法，属于光谱分析和医学图像处理领域。本发明根据光谱和内窥镜的联合探头同时采集到的光谱和图像数据构建数据库，利用光谱数据库构建神经网络癌前病变分类模型，利用图像配准和融合，将对应检测点光谱诊断结果用不同颜色显示在对应检测点图像，重构成癌前病变分布图，根据分布图输出诊断结果。本发明可以单独作为宫颈癌筛查算法，也可以为进一步活检提供图像向导。本方法的诊断结果为一张宫颈癌前病变分布图，客观直接，不受操作人员的技术影响。

