



(12)发明专利申请

(10)申请公布号 CN 110189383 A  
(43)申请公布日 2019. 08. 30

(21)申请号 201910568779.X

(22)申请日 2019.06.27

(71)申请人 合肥云诊信息科技有限公司  
地址 230000 安徽省合肥市高新区习友路  
3333号中国(合肥)国际智能语音产业  
园研发中心楼609室

(72)发明人 彭成东 王勇 杨诺 黄稳  
陈仁明 董昌武

(74)专利代理机构 合肥律众知识产权代理有限  
公司 34147

代理人 殷娟

(51) Int. Cl.  
G06T 7/90(2017.01)  
G06K 9/62(2006.01)  
G16H 30/20(2018.01)

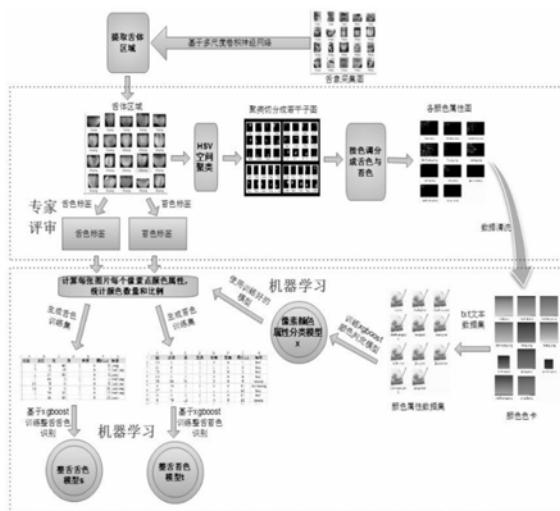
权利要求书4页 说明书12页 附图3页

(54)发明名称

基于机器学习的中医舌色苔色定量分析方法

(57)摘要

本发明涉及舌色苔色鉴定,具体涉及基于机器学习的中医舌色苔色定量分析方法,对不同地区、年龄的人群进行舌象调查与图像采集,并通过一致评审的方法对采集的舌象图像进行舌色、苔色判定,构建中医舌象大数据,对中医舌象大数据中的舌象图像提取舌体区域,通过HSV空间聚类,切分成子图,聚类结果按色调自动生成舌色颜色色卡、苔色颜色色卡,基于xgboost机器学习算法构建像素颜色属性分类器X,基于xgboost机器学习算法分别构建整舌舌色分类模型s和整舌苔色分类模型t;本发明提供的技术方案能够有效克服现有技术所存在的识别结果准确性较低、较为依赖样本数据的缺陷。



1. 基于机器学习的中医舌色苔色定量分析方法,其特征在于:包括以下步骤:

S1、对不同地区、年龄的人群进行舌象调查与图像采集,并通过一致评审的方法对采集的舌象图像进行舌色、苔色判定,构建中医舌象大数据;

S2、对中医舌象大数据中的舌象图像提取舌体区域,通过HSV空间聚类,切分成子图,聚类结果按色调自动生成舌色颜色色卡、苔色颜色色卡;

S3、基于xgboost机器学习算法构建像素颜色属性分类器X,基于xgboost机器学习算法分别构建整舌舌色分类模型s和整舌苔色分类模型t;

S4、使用像素颜色属性分类器X计算舌体区域每个像素点的颜色属性,得到颜色数量c、颜色比例f;

S5、使用整舌舌色分类模型s计算整舌舌色,使用整舌苔色分类模型t计算整舌苔色;

S6、按局部特征定义要求,对舌体各分区进行局部舌色、苔色特征计算。

2. 根据权利要求1所述的基于机器学习的中医舌色苔色定量分析方法,其特征在於:所述舌象图像为在自然光条件下,采用图像采集装置对儿童、青少年、中老年各年龄段人群进行采集。

3. 根据权利要求1所述的基于机器学习的中医舌色苔色定量分析方法,其特征在於:所述中医舌象大数据包括舌色、舌形、苔质、苔色,所述舌色包括淡白、淡红、红、绛、青紫,所述舌形包括胖大、瘦小、齿痕、裂纹、瘀斑、瘀点、点刺,所述苔质包括厚苔、薄苔、腐苔、腻苔、润苔、燥苔、剥苔,所述苔色包括白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔。

4. 根据权利要求1所述的基于机器学习的中医舌色苔色定量分析方法,其特征在於:从所述舌象图像提取舌体区域基于多尺度卷积神经网络的像素级实例分割模型,所述实例分割模型通过深度学习对舌象图像提取舌体区域并计算轮廓边界,每张舌体区域图像生成一张背景为黑色、前景为舌体的图像,作为舌体颜色特征计算数据源。

5. 根据权利要求1所述的基于机器学习的中医舌色苔色定量分析方法,其特征在於:通过所述HSV空间聚类生成舌色颜色色卡、苔色颜色色卡包括以下步骤:

S1、通过颜色聚类的方式将提取的舌体区域图像进行HSV颜色空间聚类,按照相邻区域相近颜色为同一区域切分成子图;

S2、对切分成的子图,使用抠图工具按照舌色、苔色的顺序,新建11张各颜色命名的黑色背景图片,从子图中抠出可能具有的以上颜色区域小图放置到对应的黑色背景图片中,经过反复取小图的方式,可以得到比较可靠的各颜色小块图片的拼图;

S3、对各颜色小块图片进行像素点合并,并对色调跳跃大的以及重复像素点进行数据清洗,最后生成各种颜色对应的像素点图片,即五张舌色颜色色卡、六张苔色颜色色卡。

6. 根据权利要求1所述的基于机器学习的中医舌色苔色定量分析方法,其特征在於:构建所述像素颜色属性分类器X包括将舌色颜色色卡、苔色颜色色卡像素点的HSV三个值作为三列训练集特征数据,色卡对应的颜色作为标签值,并用数字代替,制作好的数据进行颜色属性分类器X的训练,得到所述颜色属性分类器X。

7. 根据权利要求1所述的基于机器学习的中医舌色苔色定量分析方法,其特征在於:构建所述整舌舌色分类模型s和整舌苔色分类模型t包括以下步骤:

S1、使用颜色属性分类器X计算舌体区域每个像素点的颜色属性,得到颜色数量c、颜色比例f;

S2、舌色训练数据列包括淡白、淡红、红、绛、青紫颜色像素点占比及中医专家评审舌色标签,苔色数据列包括白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔颜色像素点占比及中医专家评审苔色标签;

S3、使用SVM模型对整舌舌色预测模型、整舌苔色预测模型进行数据训练,得到整舌舌色分类模型s、整舌苔色分类模型t。

8.根据权利要求1所述的基于机器学习的中医舌色苔色定量分析方法,其特征在于:使用所述整舌舌色分类模型s计算整舌舌色,使用所述整舌苔色分类模型t计算整舌苔色包括以下步骤:

S1、读取r张舌象图像,图片宽度为w,高度为h,计算各像素在HSV空间的矩阵M:

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{1w} \\ m_{21} & m_{22} & m_{2w} \\ \vdots & \vdots & \vdots \\ m_{h1} & m_{h2} & m_{hw} \end{bmatrix}_r$$

其中, $m_{ij}=[h \ s \ v]$ , $h \in [0,180]$  $s \in [0,255]$  $v \in [0,255]$

S2、使用像素颜色属性分类器X计算舌象图像中各像素属于淡白、淡红、红、绛、青紫的颜色属性类别矩阵 $N_s$ :

$$N_s = xgboost_x(M) = \begin{bmatrix} X_{11} & X_{12} & X_{1w} \\ X_{21} & X_{22} & X_{2w} \\ \vdots & \vdots & \vdots \\ X_{h1} & X_{h2} & X_{hw} \end{bmatrix}_r$$

S3、对颜色属性类别矩阵 $N_s$ 统计各类别像素数量矩阵 $C_s$ 及其占整舌比例 $f_s$ :

$$C_s = count(N_s) = \begin{bmatrix} c_{11} & c_{12} & c_{15} \\ c_{21} & c_{22} & c_{25} \\ \vdots & \vdots & \vdots \\ c_{r1} & c_{r2} & c_{r5} \end{bmatrix}$$

$$f_s = \frac{C_s}{w * h} = \begin{bmatrix} f_{11} & f_{12} & f_{15} \\ f_{21} & f_{22} & f_{25} \\ \vdots & \vdots & \vdots \\ f_{r1} & f_{r2} & f_{r5} \end{bmatrix}$$

S4、使用整舌舌色分类模型s计算整舌舌色属于淡白、淡红、红、绛、青紫的概率 $p_s$ :

$$p_s = xgboost_s(f_s) = \begin{bmatrix} p_{11} & p_{12} & p_{15} \\ p_{21} & p_{22} & p_{25} \\ p_{r1} & p_{r2} & p_{r5} \end{bmatrix}$$

取颜色最大概率 $\max(p_s) = [p_j]_r$ 为整舌舌色；

S5、使用像素颜色属性分类器X计算舌象图像中各像素属于白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔的颜色属性类别矩阵 $N_t$ ：

$$N_t = xgboost_x(M) = \begin{bmatrix} X_{11} & X_{12} & X_{1w} \\ X_{21} & X_{22} & X_{2w} \\ X_{h1} & X_{h2} & X_{hw} \end{bmatrix}_r$$

S6、对颜色属性类别矩阵 $N_t$ 统计各类别像素数量矩阵 $C_t$ 及其占整舌比例 $f_t$ ：

$$c_t = count(N_t) = \begin{bmatrix} c_{11} & c_{12} & c_{16} \\ c_{21} & c_{22} & c_{26} \\ c_{r1} & c_{r2} & c_{r6} \end{bmatrix}$$

$$f_t = \frac{c_t}{w * h} = \begin{bmatrix} f_{11} & f_{12} & f_{16} \\ f_{21} & f_{22} & f_{26} \\ f_{r1} & f_{r2} & f_{r6} \end{bmatrix}$$

S7、使用整舌苔色分类模型t计算整舌苔色属于白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔的概率 $p_t$ ：

$$p_t = xgboost_t(f_t) = \begin{bmatrix} p_{11} & p_{12} & p_{16} \\ p_{21} & p_{22} & p_{26} \\ p_{r1} & p_{r2} & p_{r6} \end{bmatrix}$$

取颜色最大概率 $\max(p_t) = [p_j]_r$ 为整舌苔色。

9. 根据权利要求1所述的基于机器学习的中医舌色苔色定量分析方法，其特征在于：所

述舌体各分区包括舌尖、舌左边、舌右边、舌中、舌根。

## 基于机器学习的中医舌色苔色定量分析方法

### 技术领域

[0001] 本发明涉及舌色苔色鉴定,具体涉及基于机器学习的中医舌色苔色定量分析方法。

### 背景技术

[0002] 传统望舌是以肉眼观察为主,其正确性取决于医生经验,受限于环境因素,缺乏客观统一的辨识指标,无法达到研究上的可重复性要求。医生可能会因光线、环境等因素,造成对舌色、苔色的判断失误,因此急需采用机器量化分析和辨识的方法。现有技术主要有以下两种方法:

[0003] 卷积神经网络分类模型

[0004] 直接使用神经网络来分类判别舌体苔质颜色,以InceptionV4分类网络为例,该网络在ILSVRC-2012-CLS数据集下以最大概率的类别作为预测结果,准确率为80.2%,而使用的是包含百万张图片的数据集。对于中医舌诊来说,要使用权威医生标注的高质量样本不仅匮乏而且代价十分昂贵,经验不足的医生往往难以保证数据标注的质量,所以能够作为训练的舌象库样本容量太小,无法达到预期的分类效果。

[0005] 卷积网络滑动窗口预测

[0006] 在《一种基于卷积神经网络的中医舌色苔色自动分析方法》专利中,提到的不再使用原图进行直接训练,而是利用滑动窗口从舌象的舌质、舌苔区域得到小块图片,然后再由中医进行确定标签,随后训练,预测的时候也是先通过滑动窗口将图片分成小块,然后通过模型预测每块的类型,综合起来判定整舌的舌色、苔色。上述方式较第一种确实有所改进,但是其中也有问题,比如使用滑动窗口方式取分块,具有随机性和不稳定性,往往一块图片上带有其他很多颜色,导致标记出现偏差,致使最终结果不准确。

### 发明内容

[0007] (一)解决的技术问题

[0008] 针对现有技术所存在的上述缺点,本发明提供了基于机器学习的中医舌色苔色定量分析方法,能够有效克服现有技术所存在的识别结果准确性较低、较为依赖样本数据的缺陷。

[0009] (二)技术方案

[0010] 为实现以上目的,本发明通过以下技术方案予以实现:

[0011] 基于机器学习的中医舌色苔色定量分析方法,包括以下步骤:

[0012] S1、对不同地区、年龄的人群进行舌象调查与图像采集,并通过一致评审的方法对采集的舌象图像进行舌色、苔色判定,构建中医舌象大数据;

[0013] S2、对中医舌象大数据中的舌象图像提取舌体区域,通过HSV空间聚类,切分成子图,聚类结果按色调自动生成舌色颜色色卡、苔色颜色色卡;

[0014] S3、基于xgboost机器学习算法构建像素颜色属性分类器X,基于xgboost机器学习

算法分别构建整舌舌色分类模型s和整舌苔色分类模型t;

[0015] S4、使用像素颜色属性分类器X计算舌体区域每个像素点的颜色属性,得到颜色数量c、颜色比例f;

[0016] S5、使用整舌舌色分类模型s计算整舌舌色,使用整舌苔色分类模型t计算整舌苔色;

[0017] S6、按局部特征定义要求,对舌体各分区进行局部舌色、苔色特征计算。

[0018] 优选地,所述舌象图像为在自然光条件下,采用图像采集装置对儿童、青少年、中老年各年龄段人群进行采集。

[0019] 优选地,所述中医舌象大数据包括舌色、舌形、苔质、苔色,所述舌色包括淡白、淡红、红、绛、青紫,所述舌形包括胖大、瘦小、齿痕、裂纹、瘀斑、瘀点、点刺,所述苔质包括厚苔、薄苔、腐苔、腻苔、润苔、燥苔、剥苔,所述苔色包括白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔。

[0020] 优选地,从所述舌象图像提取舌体区域基于多尺度卷积神经网络的像素级实例分割模型,所述实例分割模型通过深度学习对舌象图像提取舌体区域并计算轮廓边界,每张舌体区域图像生成一张背景为黑色、前景为舌体的图像,作为舌体颜色特征计算数据源。

[0021] 优选地,通过所述HSV空间聚类生成舌色颜色色卡、苔色颜色色卡包括以下步骤:

[0022] S1、通过颜色聚类的方式将提取的舌体区域图像进行HSV颜色空间聚类,按照相邻区域相近颜色为同一区域切分成子图;

[0023] S2、对切分成的子图,使用抠图工具按照舌色、苔色的顺序,新建11张各颜色命名的黑色背景图片,从子图中抠出可能具有的以上颜色区域小图放置到对应的黑色背景图片中,经过反复取小图的方式,可以得到比较可靠的各颜色小块图片的拼图;

[0024] S3、对各颜色小块图片进行像素点合并,并对色调跳跃大的以及重复像素点进行数据清洗,最后生成各种颜色对应的像素点图片,即五张舌色颜色色卡、六张苔色颜色色卡。

[0025] 优选地,构建所述像素颜色属性分类器X包括将舌色颜色色卡、苔色颜色色卡像素点的HSV三个值作为三列训练集特征数据,色卡对应的颜色作为标签值,并用数字代替,制作好的数据进行颜色属性分类器X的训练,得到所述颜色属性分类器X。

[0026] 优选地,构建所述整舌舌色分类模型s和整舌苔色分类模型t包括以下步骤:

[0027] S1、使用颜色属性分类器X计算舌体区域每个像素点的颜色属性,得到颜色数量c、颜色比例f;

[0028] S2、舌色训练数据列包括淡白、淡红、红、绛、青紫颜色像素点占比及中医专家评审舌色标签,苔色数据列包括白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔颜色像素点占比及中医专家评审苔色标签;

[0029] S3、使用SVM模型对整舌舌色预测模型、整舌苔色预测模型进行数据训练,得到整舌舌色分类模型s、整舌苔色分类模型t。

[0030] 优选地,使用所述整舌舌色分类模型s计算整舌舌色,使用所述整舌苔色分类模型t计算整舌苔色包括以下步骤:

[0031] S1、读取r张舌象图像,图片宽度为w,高度为h,计算各像素在HSV空间的矩阵M:

$$[0032] \quad M = \begin{bmatrix} m_{11} & m_{12} & m_{1w} \\ m_{21} & m_{22} & m_{2w} \\ \vdots & \vdots & \vdots \\ m_{h1} & m_{h2} & m_{hw} \end{bmatrix}_r$$

[0033] 其中,  $m_{ij} = [h \ s \ v]$ ,  $h \in [0, 180]$   $s \in [0, 255]$   $v \in [0, 255]$

[0034] S2、使用像素颜色属性分类器X计算舌象图像中各像素属于淡白、淡红、红、绛、青紫的颜色属性类别矩阵 $N_s$ :

$$[0035] \quad N_s = xgboost_x(M) = \begin{bmatrix} X_{11} & X_{12} & X_{1w} \\ X_{21} & X_{22} & X_{2w} \\ \vdots & \vdots & \vdots \\ X_{h1} & X_{h2} & X_{hw} \end{bmatrix}_r$$

[0036] S3、对颜色属性类别矩阵 $N_s$ 统计各类别像素数量矩阵 $C_s$ 及其占整舌比例 $f_s$ :

$$[0037] \quad c_s = count(N_s) = \begin{bmatrix} c_{11} & c_{12} & c_{15} \\ c_{21} & c_{22} & c_{25} \\ \vdots & \vdots & \vdots \\ c_{r1} & c_{r2} & c_{r5} \end{bmatrix}$$

$$[0038] \quad f_s = \frac{c_s}{w * h} = \begin{bmatrix} f_{11} & f_{12} & f_{15} \\ f_{21} & f_{22} & f_{25} \\ \vdots & \vdots & \vdots \\ f_{r1} & f_{r2} & f_{r5} \end{bmatrix}$$

[0039] S4、使用整舌舌色分类模型s计算整舌舌色属于淡白、淡红、红、绛、青紫的概率 $p_s$ :

$$[0040] \quad p_s = xgboost_s(f_s) = \begin{bmatrix} p_{11} & p_{12} & p_{15} \\ p_{21} & p_{22} & p_{25} \\ \vdots & \vdots & \vdots \\ p_{r1} & p_{r2} & p_{r5} \end{bmatrix}$$

[0041] 取颜色最大概率 $\max(p_s) = [p_j]_r$ 为整舌舌色;

[0042] S5、使用像素颜色属性分类器X计算舌象图像中各像素属于白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔的颜色属性类别矩阵 $N_t$ :



$$[0043] \quad N_t = xgboost_x(M) = \begin{bmatrix} X_{11} & X_{12} & X_{1w} \\ X_{21} & X_{22} & X_{2w} \\ X_{h1} & X_{h2} & X_{hw} \end{bmatrix}_r$$

[0044] S6、对颜色属性类别矩阵 $N_t$ 统计各类别像素数量矩阵 $C_t$ 及其占整舌比例 $f_t$ ：

$$[0045] \quad c_t = count(N_t) = \begin{bmatrix} c_{11} & c_{12} & c_{16} \\ c_{21} & c_{22} & c_{26} \\ c_{r1} & c_{r2} & c_{r6} \end{bmatrix}$$

$$[0046] \quad f_t = \frac{c_t}{w * h} = \begin{bmatrix} f_{11} & f_{12} & f_{16} \\ f_{21} & f_{22} & f_{26} \\ f_{r1} & f_{r2} & f_{r6} \end{bmatrix}$$

[0047] S7、使用整舌苔色分类模型 $t$ 计算整舌苔色属于白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔的概率 $p_t$ ：

$$[0048] \quad p_t = xgboost_t(f_t) = \begin{bmatrix} p_{11} & p_{12} & p_{16} \\ p_{21} & p_{22} & p_{26} \\ p_{r1} & p_{r2} & p_{r6} \end{bmatrix}$$

[0049] 取颜色最大概率 $\max(p_t) = [p_j]_r$ 为整舌苔色。

[0050] 优选地，所述舌体各分区包括舌尖、舌左边、舌右边、舌中、舌根。

[0051] (三) 有益效果

[0052] 与现有技术相比，本发明所提供的基于机器学习的中医舌色苔色定量分析方法具有以下有益效果：

[0053] (1)、准确度明显高于同类方法，对舌体颜色特征的检测能力高于一般方法，检测算法的准确性更高，所需数据依赖更小，更能保证日常自然光照条件下拍照舌诊使用；

[0054] (2)、通过对HSV颜色空间舌色、苔色属性的两次机器学习，对舌体进行像素级分类和舌色、苔色的定量计算，能达到对舌体5种舌色、6种苔色，共11个类别的检测，以及舌尖、舌中、舌根、舌左边、舌右边5个区域局部颜色的特征识别，显著提升了检测能力并且重复性好、准确度高，可应用于人工智能舌诊机器辨析舌色、苔色。

## 附图说明

[0055] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术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发明的一些实施例,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。

[0056] 图1为本发明舌色、苔色机器学习模型训练流程示意图;

[0057] 图2为本发明舌色、苔色分类模型定量分析示意图;

[0058] 图3为本发明舌体区域划分示意图。

## 具体实施方式

[0059] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述。显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0060] 基于机器学习的中医舌色苔色定量分析方法,如图1至图3所示,包括以下步骤:

[0061] S1、对不同地区、年龄的人群进行舌象调查与图像采集,并通过一致评审的方法对采集的舌象图像进行舌色、苔色判定,构建中医舌象大数据;

[0062] S2、对中医舌象大数据中的舌象图像提取舌体区域,通过HSV空间聚类,切分成子图,聚类结果按色调自动生成舌色颜色色卡、苔色颜色色卡;

[0063] S3、基于xgboost机器学习算法构建像素颜色属性分类器X,基于xgboost机器学习算法分别构建整舌舌色分类模型s和整舌苔色分类模型t;

[0064] S4、使用像素颜色属性分类器X计算舌体区域每个像素点的颜色属性,得到颜色数量c、颜色比例f;

[0065] S5、使用整舌舌色分类模型s计算整舌舌色,使用整舌苔色分类模型t计算整舌苔色;

[0066] S6、按局部特征定义要求,对舌体各分区进行局部舌色、苔色特征计算。

[0067] 舌象图像为在自然光条件下,采用图像采集装置对儿童、青少年、中老年各年龄段人群进行采集。

[0068] 中医舌象大数据包括舌色、舌形、苔质、苔色,舌色包括淡白、淡红、红、绛、青紫,舌形包括胖大、瘦小、齿痕、裂纹、瘀斑、瘀点、点刺,苔质包括厚苔、薄苔、腐苔、腻苔、润苔、燥苔、剥苔,苔色包括白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔。

[0069] 从舌象图像提取舌体区域基于多尺度卷积神经网络的像素级实例分割模型,实例分割模型通过深度学习对舌象图像提取舌体区域并计算轮廓边界,每张舌体区域图像生成一张背景为黑色、前景为舌体的图像,作为舌体颜色特征计算数据源。

[0070] 通过HSV空间聚类生成舌色颜色色卡、苔色颜色色卡包括以下步骤:

[0071] S1、通过颜色聚类的方式将提取的舌体区域图像进行HSV颜色空间聚类,按照相邻区域相近颜色为同一区域切分成子图;

[0072] S2、对切分成的子图,使用抠图工具按照舌色、苔色的顺序,新建11张各颜色命名的黑色背景图片,从子图中抠出可能具有的以上颜色区域小图放置到对应的黑色背景图片

中,经过反复取小图的方式,可以得到比较可靠的各颜色小块图片的拼图;

[0073] S3、对各颜色小块图片进行像素点合并,并对色调跳跃大的以及重复像素点进行数据清洗,最后生成各种颜色对应的像素点图片,即五张舌色颜色色卡、六张苔色颜色色卡。

[0074] 构建像素颜色属性分类器X包括将舌色颜色色卡、苔色颜色色卡像素点的HSV三个值作为三列训练集特征数据,色卡对应的颜色作为标签值,并用数字代替,制作好的数据进行颜色属性分类器X的训练,得到颜色属性分类器X。

[0075] 构建整舌舌色分类模型s和整舌苔色分类模型t包括以下步骤:

[0076] S1、使用颜色属性分类器X计算舌体区域每个像素点的颜色属性,得到颜色数量c、颜色比例f;

[0077] S2、舌色训练数据列包括淡白、淡红、红、绛、青紫颜色像素点占比及中医专家评审舌色标签,苔色数据列包括白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔颜色像素点占比及中医专家评审苔色标签;

[0078] S3、使用SVM模型对整舌舌色预测模型、整舌苔色预测模型进行数据训练,得到整舌舌色分类模型s、整舌苔色分类模型t。

[0079] 使用整舌舌色分类模型s计算整舌舌色,使用整舌苔色分类模型t计算整舌苔色包括以下步骤:

[0080] S1、读取r张舌象图像,图片宽度为w,高度为h,计算各像素在HSV空间的矩阵M:

$$[0081] \quad M = \begin{bmatrix} m_{11} & m_{12} & m_{1w} \\ m_{21} & m_{22} & m_{2w} \\ \vdots & \vdots & \vdots \\ m_{h1} & m_{h2} & m_{hw} \end{bmatrix}_r$$

[0082] 其中, $m_{ij} = [h \ s \ v]$ ,  $h \in [0, 180]$   $s \in [0, 255]$   $v \in [0, 255]$

[0083] S2、使用像素颜色属性分类器X计算舌象图像中各像素属于淡白、淡红、红、绛、青紫的颜色属性类别矩阵 $N_s$ :

$$[0084] \quad N_s = xgboost_x(M) = \begin{bmatrix} X_{11} & X_{12} & X_{1w} \\ X_{21} & X_{22} & X_{2w} \\ \vdots & \vdots & \vdots \\ X_{h1} & X_{h2} & X_{hw} \end{bmatrix}_r$$

[0085] S3、对颜色属性类别矩阵 $N_s$ 统计各类别像素数量矩阵 $C_s$ 及其占整舌比例 $f_s$ :

$$[0086] \quad c_s = \text{count}(N_s) = \begin{bmatrix} c_{11} & c_{12} & c_{15} \\ c_{21} & c_{22} & c_{25} \\ c_{r1} & c_{r2} & c_{r5} \end{bmatrix}$$

$$[0087] \quad f_s = \frac{c_s}{w * h} = \begin{bmatrix} f_{11} & f_{12} & f_{15} \\ f_{21} & f_{22} & f_{25} \\ f_{r1} & f_{r2} & f_{r5} \end{bmatrix}$$

[0088] S4、使用整舌舌色分类模型s计算整舌舌色属于淡白、淡红、红、绛、青紫的概率 $p_s$ ：

$$[0089] \quad p_s = \text{xgboost}_s(f_s) = \begin{bmatrix} p_{11} & p_{12} & p_{15} \\ p_{21} & p_{22} & p_{25} \\ p_{r1} & p_{r2} & p_{r5} \end{bmatrix}$$

[0090] 取颜色最大概率 $\max(p_s) = [p_j]_r$ 为整舌舌色；

[0091] S5、使用像素颜色属性分类器X计算舌象图像中各像素属于白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔的颜色属性类别矩阵 $N_t$ ：

$$[0092] \quad N_t = \text{xgboost}_x(M) = \begin{bmatrix} X_{11} & X_{12} & X_{1w} \\ X_{21} & X_{22} & X_{2w} \\ X_{h1} & X_{h2} & X_{hw} \end{bmatrix}_r$$

[0093] S6、对颜色属性类别矩阵 $N_t$ 统计各类别像素数量矩阵 $C_t$ 及其占整舌比例 $f_t$ ：

$$[0094] \quad c_t = \text{count}(N_t) = \begin{bmatrix} c_{11} & c_{12} & c_{16} \\ c_{21} & c_{22} & c_{26} \\ c_{r1} & c_{r2} & c_{r6} \end{bmatrix}$$

$$[0095] \quad f_t = \frac{c_t}{w * h} = \begin{bmatrix} f_{11} & f_{12} & f_{16} \\ f_{21} & f_{22} & f_{26} \\ f_{r1} & f_{r2} & f_{r6} \end{bmatrix}$$

[0096] S7、使用整舌苔色分类模型t计算整舌苔色属于白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔的概率 $p_t$ :

$$[0097] \quad p_t = xgboost_t(f_t) = \begin{bmatrix} p_{11} & p_{12} & p_{16} \\ p_{21} & p_{22} & p_{26} \\ p_{r1} & p_{r2} & p_{r6} \end{bmatrix}$$

[0098] 取颜色最大概率 $\max(p_t) = [p_j]_r$ 为整舌苔色。

[0099] 舌体各分区包括舌尖、舌左边、舌右边、舌中、舌根。

[0100] 如图3所示,舌体分区是将舌体划分成左、右、上、下、中5个部分。取舌体的最小外接矩形,将左右两边的上1/5处进行连线,将左右两边的下1/5处进行连线,上下两边的左1/5处进行连线,上下两边的右1/5处进行连线,通过4条线将最小外接矩形分成了A、B、C、D、E、F、G、H、I共9块区域,将B区域记为舌根,A、D、G区域记为舌左边,C、F、I区域记为舌右边、E区域记舌中、H区域记为舌尖。

[0101] 使用各颜色分类数量 $C_s$ 、 $C_t$ 得到各所属区域颜色分类数量,根据各类颜色区域像素量与舌体的区域像素重合比例,以舌体五个区域中重合比例最高的记为舌色、苔色的主要分布区域。局部特征定义和阈值化计算公式如表1:

[0102] 表1局部特征定义和阈值化计算公式

[0103]

局部特征	特征定义	阈值化计算公式
舌尖红	舌尖区域超过一半有红绛色时, 则判定舌象 $T_i$ 为舌尖红	$Tr_i = \frac{c_{si3} + c_{si4}}{H} \geq 50\%$
舌两边红	舌左边和右边区域超过一半有红绛色时, 则判定舌象 $T_i$ 为舌两边红	$Tr_i = \frac{c_{si3} + c_{si4}}{D+F} \geq 50\%$
舌中厚白苔	舌中区域超过一半有厚白苔时, 则判定舌象 $T_i$ 为舌中厚白苔	$Tw_i = \frac{c_{ti1}}{E} \geq 50\%$
舌根厚白苔	舌根区域超过一半有厚白苔时, 则判定舌象 $T_i$ 为舌中厚白苔	$Tw_i = \frac{c_{ti1}}{B} \geq 50\%$
舌中黄苔	舌中区域超过一半有淡黄、黄、焦黄苔时, 则判定舌象 $T_i$ 为舌中黄苔	$Ty_i = \frac{c_{ti2} + c_{ti3} + c_{ti4}}{E} \geq 50\%$
舌根黄苔	舌根区域超过一半有淡黄、黄、焦黄苔时, 则判定舌象 $T_i$ 为舌根黄苔	$Ty_i = \frac{c_{ti2} + c_{ti3} + c_{ti4}}{B} \geq 50\%$
舌苔满布	整舌舌苔超 90%以上时为苔满布	$Tf_i = \frac{\sum_{j=1}^6 c_{tij}}{B} \geq 90\%$
无苔或假苔	整舌苔低于 5%时为无苔或假苔	$Te_i = \frac{\sum_{j=1}^6 c_{tij}}{B} \leq 5\%$

[0104] 从1000张测试集舌象图像中, 随机挑选舌象图像中若干像素点, 选用不重复2000个像素点, 将每个像素点扩大成40\*40的小块图片, 从中选取550张样本图, 由中医专家和像素颜色属性分类器X同时进行颜色辨识, 使用像素颜色属性分类器X对颜色分类的准确率高于85%, 实验结果分析如表2:

[0105] 表2像素点颜色预测结果

[0106]

数据集		中医专家辨识(张)	分类器 X 正确数(张)	准确率
舌色	淡红	72	72	100%
	淡白	56	56	98.6%
	红	51	49	96.1%
	绛	44	42	95.5%
	青紫	15	13	85.0%
苔色	白	90	89	98.9%
	淡黄	67	65	97.0%
	黄	73	70	95.9%
	焦黄	42	41	97.6%
	灰黑	24	21	87.5%
	焦黑	16	14	87.5%
统计	-	550	530	96.4%

[0107] 结果显示:本发明训练得到像素颜色属性分类器X的准确性与专家评审的一致率保持较高水平,具备对舌色、苔色的辨识能力。

[0108] 从中医舌象大数据中抽取5000例舌象图像,按舌色分类建立淡白舌、淡红舌、红舌、绛舌、青紫舌数据集,利用神经网络、图像属性阈值算法、两次机器学习三种算法进行实验,使用xgboost两次机器学习预测准确率达85%以上,舌色实验结果分析如表3:

[0109] 表3舌色颜色属性及综合判定分析准确率结果

[0110]

舌象数据集	神经网络	机器学习+阈值算法	两次机器学习
淡白舌	70%	60%	90%
淡舌	70%	70%	85%
淡红舌	70%	70%	85%
红舌	70%	70%	90%
绛舌	70%	70%	90%
青紫舌	70%	70%	90%

[0111] 结果显示:本发明在经过两次机器学习训练后,对整舌舌色判定的准确性更高。

[0112] 对白苔、淡黄苔、黄苔、焦黄苔、灰黑苔、焦黑苔共六类数据集实验分析,利用神经网络、图像属性阈值算法、两次机器学习三种算法进行实验,使用xgboost两次机器学习预测的准确率达85%,苔色实验结果分析如表4:

[0113] 表4苔色颜色属性及综合判定分析准确率结果

[0114]

舌象数据集	神经网络	机器学习+阈值算法	两次机器学习
白苔舌	70%	60%	85%
淡黄苔舌	70%	70%	85%
黄苔舌	70%	70%	90%
焦黄苔舌	70%	70%	85%
灰黑苔舌	70%	70%	90%
焦黑苔舌	70%	70%	85%

[0115] 结果显示:本发明在经过两次机器学习训练后,对整舌苔色判定的准确性更高。

[0116] 从中医舌象大数据中抽取具有舌尖红图片200张、无舌尖红图片200张、舌两边红图片200张、舌两边不红图片200张,按舌尖红和两边红组建数据集,进行舌尖红和舌两边红判定实验。在数据量较少的情况下,用神经网络对舌尖红、舌两边红分类准确率低,用机器学习与阈值算法判定准确率在85%,在舌尖不红的情况下误检测概率小。舌色局部特征实验结果见下表5:

[0117] 表5舌色局部特征的检测结果

[0118]

舌象数据集	神经网络	机器学习+阈值算法
舌尖红	75%	90%
舌尖不红	20%	85%



舌两边红	75%	90%
舌两边不红	20%	85%

[0119] 结果显示:本发明方法对舌尖红、舌两边红特征的进一步判定,准确性高,出现误检的概率低。

[0120] 对中根厚白苔、中根无厚白苔、中根黄苔、中根无黄苔、苔满布、无苔各200张舌象图像,共六类数据集实验分析。在数据量较少情况下,用神经网络对舌尖红、舌两边红分类准确率低,使用机器学习与阈值算法判定准确率在85%左右。舌苔局部特征实验结果分析如表6:

[0121] 表6舌苔局部特征的检测结果

[0122]

舌象数据集	神经网络	机器学习+阈值算法
中根厚白苔	75%	85%
中根无厚白苔	20%	85%
中根黄苔	75%	90%
中根无黄苔	20%	85%
舌苔满布	70%	90%
无苔	70%	80%

[0123] 结果显示:本发明对舌中根厚白苔、中根黄苔局部特征的进一步判定,准确性较高,出现误检的概率低。

[0124] 以上实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不会使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围。

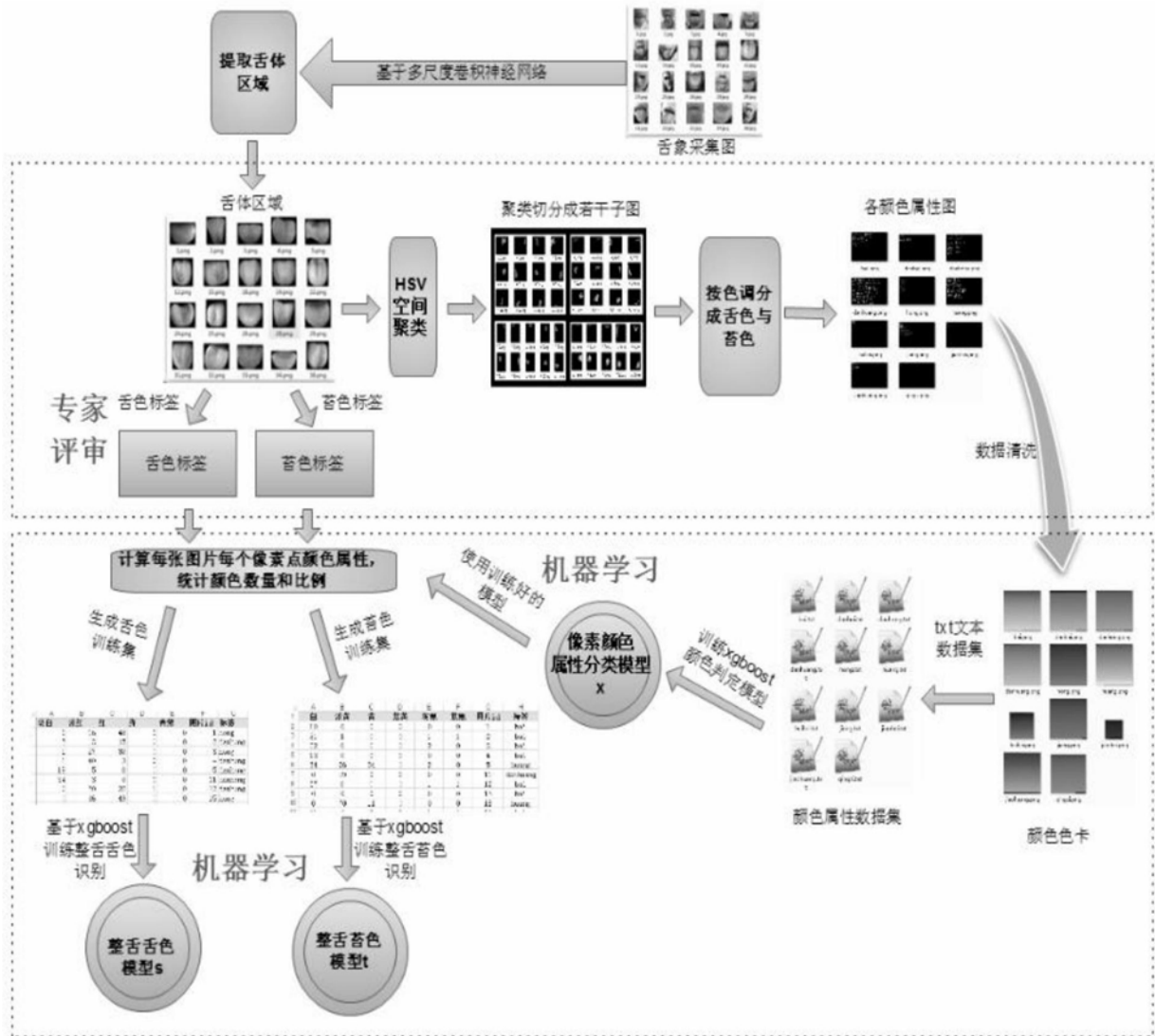


图1

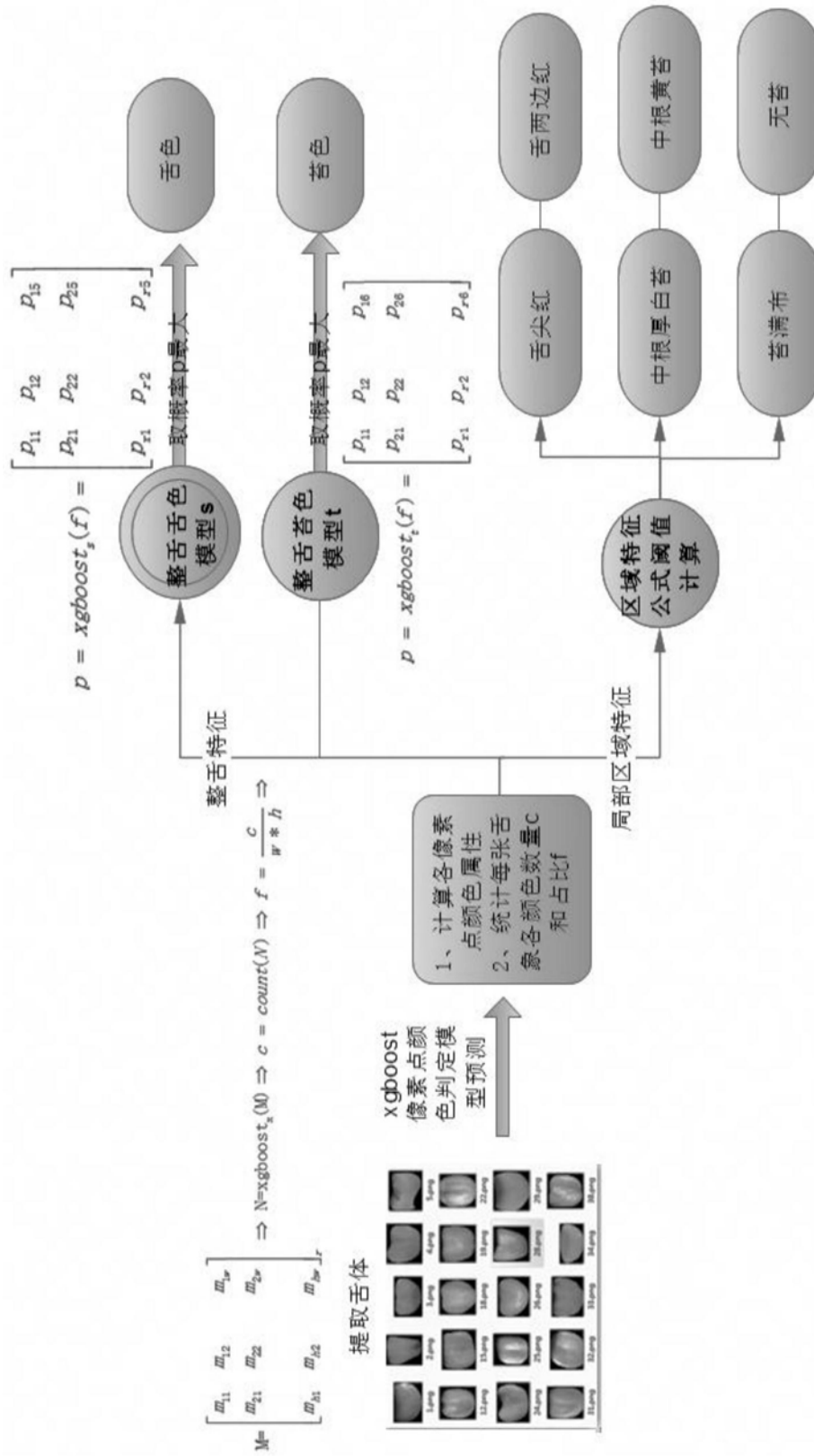


图2

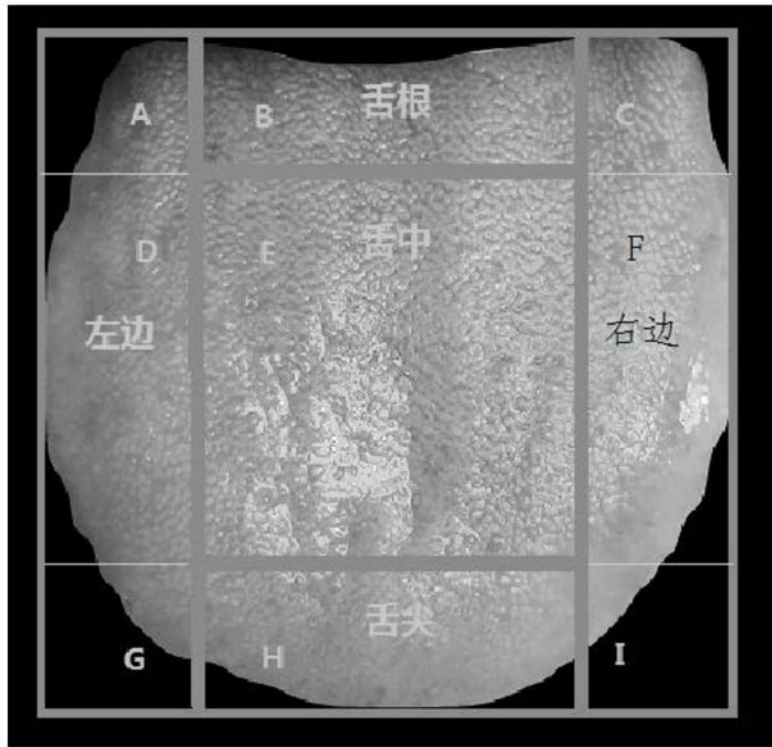


图3