

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
A61B 5/05 (2006.01)



# [12] 发明专利申请公布说明书

[21] 申请号 200480027236.9

[43] 公开日 2007年3月14日

[11] 公开号 CN 1929781A

[22] 申请日 2004.8.21  
 [21] 申请号 200480027236.9  
 [30] 优先权  
     [32] 2003.8.21 [33] US [31] 60/497,375  
 [86] 国际申请 PCT/US2004/027171 2004.8.21  
 [87] 国际公布 WO2005/020790 英 2005.3.10  
 [85] 进入国家阶段日期 2006.3.21  
 [71] 申请人 依斯克姆公司  
     地址 美国加利福尼亚州  
 [72] 发明人 乔尔·T·休伊曾加  
             鲁塞尔·W·安德森  
             托马斯·伍德利·布拉泽顿

[74] 专利代理机构 北京康信知识产权代理有限责任  
                     公司  
                     代理人 章社杲

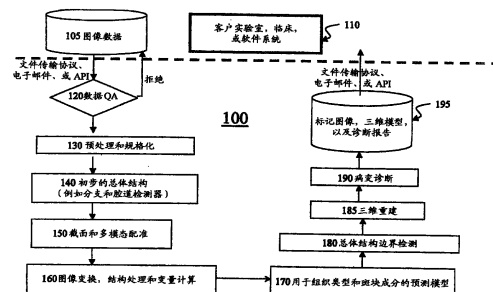
权利要求书 7 页 说明书 48 页 附图 14 页

## [54] 发明名称

用于脉管斑块检测和自动化方法和系统

## [57] 摘要

用于检测和分析患者脉管系统的一个或更多区域中的斑块的自动化方法和系统。



1. 一种评定在患者脉管系统的至少一部分中动脉粥样硬化程度的自动化方法,其包含计算处理来自于医学成像系统的患者脉管系统的至少一个血管的至少一个横截面的可处理数据,以确定是否所述血管包含至少一个和斑块的存在相关的组织,从而评定所述患者脉管系统的至少一个部位中的动脉粥样硬化程度。
2. 一种确定患者脉管系统的血管是否含有斑块的自动化方法,其包括计算处理来自于医学成像系统的患者脉管系统的血管的至少一个横截面的可处理数据,以确定是否所述血管包含至少一个和斑块的存在相关的组织,在该情况下所述血管被确定为含有斑块。
3. 根据权利要求2所述的自动化方法,其中所述医学成像系统是非侵入性成像系统。
4. 根据权利要求3所述的自动化方法,其中所述非侵入性成像系统包括选自: MRI 设备、CT 设备、PET 设备、温度记录设备、以及超声设备的一种或多种设备,以生成从其衍生所述可处理数据的原始图像数据。
5. 根据权利要求2所述的自动化方法,其中所述医学成像系统包括 MRI 设备,以生成从其衍生可处理磁共振数据的原始磁共振数据。
6. 根据权利要求5所述的自动化方法,其中使用了很多不同的射频脉冲序列系列,其中每一个所述不同射频脉冲序列系列允许

为可存在于所述血管中的不同组织的子集产生原始磁共振数据。

7. 根据权利要求 6 所述的自动化方法,其中至少所述不同射频脉冲序列系列之一导致原始磁共振数据的产生,所述数据选自由 T1-加权数据, T2-加权数据, PDW-加权数据, 和 TOF-加权数据组成的组。
8. 根据权利要求 2 所述的自动化方法,其中所述血管包括供血到器官的脉管系统的部分, 所述器官选自由脑和心脏组成的组。
9. 根据权利要求 2 所述的自动化方法,其中所述血管选自由颈动脉和冠状动脉组成的组。
10. 根据权利要求 2 所述的自动化方法, 其中所述患者是人。
11. 根据权利要求 2 所述的自动化方法,其中所述可处理数据是通过预处理由所述医学成像系统生成的原始数据产生的。
12. 根据权利要求 11 所述的自动化方法, 进一步包含在计算处理所述可处理数据之前规格化所述可处理数据。
13. 根据权利要求 1 所述的自动化方法,其中计算处理所述可处理数据以确定所述血管,在所述横截面的区域,是否包含动脉和斑块。
14. 根据权利要求 13 所述的自动化方法, 其中确定组织类型是通过借助计算机比较在所述数据中鉴别的不同组织类型和许多统计分类装置来完成。

15. 根据权利要求 14 所述的自动化方法, 其中所述统计分类装置是使用已知结果数据并借助一方法而开发的, 其中所述方法选自由逻辑回归, 判定树, 非参数回归, Fisher 判别分析, Bayesian 网络建模、和模糊逻辑系统组成的组。
16. 根据权利要求 15 所述的自动化方法, 其中借助一方法测定很多储存的组织分类装置要素中的至少之一, 其中所述方法选自由术后组织学检查, 直接组织观查, 和由一个或多个专家进行标记组成的组。
17. 根据权利要求 14 所述的自动化方法, 其中计算处理所述可处理数据以确定所述血管, 在所述横截面的区域, 是否进一步包含至少一种组织, 所述组织选自由外膜, 钙沉积物, 胆固醇沉积物, 纤维斑块, 和血栓组成的组。
18. 根据权利要求 17 所述的自动化方法, 其中组织类型的确定是通过借助计算机比较在所述数据中识别的不同组织类型和许多统计分类装置来完成。
19. 根据权利要求 6 所述的自动化方法, 其中包含所述可处理磁共振数据的部分是计算上引入配准。
20. 根据权利要求 19 所述的自动化方法, 其中配准是通过对准包含所述可处理磁共振数据的部分来完成的, 所述数据是关于代表标志的表示, 所述标志选自由血管腔形心和血管分支点组成的组。
21. 根据权利要求 20 所述的自动化方法, 其中所述标记是选自由物理标记和计算标记组成的组。

22. 根据权利要求 2 所述的自动化方法,其包含计算处理所述血管的许多隔开的横截面的可处理数据。
23. 根据权利要求 22 所述的自动化方法,进一步包含通过计算提供在由所述血管的最远距离隔开的横截面所限制区域的至少一部分的所述血管的三维模型。
24. 根据权利要求 2 所述的自动化方法,进一步包含生成输出文件,所述输出文件包含来自所述计算处理的数据。
25. 根据权利要求 24 所述的自动化方法,其中所述输出文件进一步包含通过计算提供的在由所述血管的最远距离隔开的横截面所限制区域的至少一部分的所述血管的三维模型。
26. 根据权利要求 23 所述的自动化方法,进一步包含计算确定存在于所述血管的三维模型中的斑块体积。
27. 根据权利要求 23 所述的自动化方法,进一步包含计算确定存在于所述血管的三维模型中的斑块成分。
28. 根据权利要求 27 所述的自动化方法,进一步包含计算区别所述斑块是易损斑块还是稳定的斑块。
29. 一种评定治疗方案的效力的自动化方法,包括:
  - a. 利用根据权利要求 26 所述的自动化方法来确定患者体内的斑块含量;
  - b. 给所述患者提供治疗方案,所述治疗方案包括给予在所述治疗方案进程中预期能稳定或减少患者体内的斑块含量的药物;以及

- c. 在所述治疗方案过程中和/或结束时, 确定是否所述患者体内的斑块含量已经稳定或减少, 从而允许评定所述治疗方案的效力。
30. 根据权利要求 24 所述的自动化方法, 其中所述药物是研究用药物。
31. 一种对在血管中有易损斑块的患者进行治疗的方法, 包括:
- a. 利用根据权利要求 27 所述的自动化方法来确定患者血管中含有易损斑块; 以及
- b. 给所述患者提供治疗方案, 所述治疗方案用来稳定或减轻在所述治疗方案进程中所述易损斑块发生破裂的易损性, 从而对所述患者进行治疗。
32. 根据权利要求 31 所述的方法, 其中治疗方案包括给予药物, 所述药物可稳定或减轻所述斑块发生破裂的易损性。
33. 根据权利要求 32 所述的方法, 其中所述药物选自由斯特汀类、消炎药、和抗凝血药组成的组。
34. 一种计算机程序产品, 包括具有计算机可读程序代码具体化在其中的计算机可用介质, 其中所述计算机可读程序代码用于以用适合于执行所述计算机可读程序代码的计算机实施根据权利要求 1 所述的自动化方法。

35. 一种用于分析利用医学成像系统获得的患者数据以确定所述患者的血管是否含有斑块的自动化方法, 包括:
- a. 获得患者脉管系统的血管的至少一个横截面的可处理数据, 其中所述可处理数据是衍生自利用医学成像系统收集的原始数据;
  - b. 将所述可处理数据传递到用于接收和计算处理所述可处理数据的计算机, 以确定是否所述横截面的区域中的血管包括至少一个与斑块的存在相关的组织, 在该情况下, 所述血管被确定含有斑块; 以及
  - c. 利用所述计算机来计算处理所述可处理数据和确定所述血管是否含有斑块。
36. 根据权利要求 35 所述的自动化方法, 其中所述医学成像系统和所述计算机放置在不同的位置。
37. 根据权利要求 36 所述的自动化方法, 其中所述计算机位于物理上离开许多成像中心的每一个的计算中心, 其中所述每个成像中心包括能够生成原始数据的医学成像系统, 而从所述原始数据可以导出可处理数据。
38. 根据权利要求 37 所述的自动化方法, 其中至少所述成像中心之一通过远程通信链路把原始数据送到所述计算中心。
39. 根据权利要求 35 所述的自动化方法, 进一步包括把所述分析的结果传送到与用来收集所述原始数据的医学成像系统相关联的指定地址。

40. 一种用于分析患者脉管系统的血管是否含有斑块的系统,所述系统包括:

a. 适合于执行计算机可读程序代码的计算机,其中所述计算机可读程序代码具体化在根据权利要求34所述的计算机程序产品中。

b. 和所述计算机相连的计算机数据储存系统,其中所述计算机数据储存系统用于发送数据到所述计算机,从所述计算机接收数据,和/或储存由所述计算机处理的数据;以及

c. 一种用于接收由所述计算机要处理的数据或用于发送由所述计算机处理后的数据的通信接口,所述通信接口操作上连通到所述计算机。



# 用于脉管斑块检测和分析的 自动化方法和系统

## 相关申请

本专利申请要求于2003年8月21日提交的美国临时专利申请第60/497,375号的优先权，其全部内容以引用方式结合于此。

## 技术领域

本发明涉及用于医学成像数据的自动化分析的方法、软件和系统。具体地说，本发明涉及用于患者的部分或所有脉管系统内的斑块的自动化检测和分析的方法、软件、和系统。

## 背景技术

### 1. 介绍

以下说明包括可以有益于理解本发明的信息。它并不是承认：任何这样的信息是目前要求的发明的现有技术或与其有关的技术，或任何明确或隐含引用的出版物是现有技术。

### 2. 背景

动脉粥样硬化是缺血性心脏病的最常见的原因。当分别考虑时，中风是死亡的第三个主要原因，其中大多数中风是缺血性事件的结果。然而，动脉硬化是相当普通的炎症反应，而没有血栓形成的动脉粥样硬化一般是良性疾病。一些研究表明斑块的组成和狭窄

程度相比更是预测易于破裂或血栓形成的关键因素。如此易于形成血栓或高风险的斑块被认为是“易损”的斑块。

斑块的破裂由机械事件触发，但斑块的易损性起因于纤维被层的弱化、斑块间出血、以及斑块成分的软化，通常是感染以及巨噬细胞和 T-细胞浸润的结果。一般来说，富脂质的软斑块比富胶原的硬斑块更倾向于破裂。一些形态学和生理学特征与易损以及稳定的斑块有关。形态学特征提示结构弱化或损伤（变细或破裂的纤维被层、钙化、阴性重塑、血管再生、大的脂质沉积等），而生理学特征提示化学组成、活性感染、炎症反应、以及新陈代谢。许多因素是主观的或定性的，反映下述事实：并不是所有特征已经被证实是危险的决定因素。危险因素的确证要求长期的纵向临床研究、动脉内膜剥除术、或尸体解剖。

一些入侵性的方法已经用于确证易损的斑块，包括静脉内超声（IVUS）、血管镜检查、脉管内 MR、和温度记录法。由于入侵性方法把患者暴露于中风和 MI 的巨大危险中，所以这些方法不适于筛查或连续检查。最后，由于这些方法需要使用导管，所以整体脉管斑块负荷的估计必须仅从一些局部斑块沉积的检查推测。而且，由于如导管、动脉尺寸、动脉分枝等的生理约束，许多患者的脉管系统是入侵性设备不能进入的。

虽然 MRI 已经被用于鉴别斑块的形态学特征，如斑块的尺寸和纤维被层的厚度，并具有高灵敏度和特异性，但大多数描绘斑块特征的工作涉及由放射学专家进行的 CAT 或 MRI 扫描的目视检查。这是耗时的（因而是昂贵的）和易出错的过程，同时易受一些主观偏见影响，尤其是众所周知人类对同时评价多于两个或三个变量之间的统计学关系是较差的。通常的趋势是关注整体边界和局部结构（纹理，texture）。当考虑到多模态图像时，这个问题被增加了几倍，因为为了领会所有可获得的证据，分析者不得不逐个像素地评估多

达四个不同模态的局部环境。典型地，这使分析者仅关注于一种模态，此模态具有特定组织的“最佳”对照，并忽略在其它模态中的潜在的相反证据。分类准确度受到研究者之间以及甚至相同研究者随时间的可变性的影响，这使标准化的诊断检测实际上是不可能的。在绝大多数情况下，解释图像的确证仅能通过动脉内膜剥除术的组织学检查来完成。

鉴于斑块检测和分析对患者健康的重要性，显然需要改进的方法以在体内对斑块进行检测和分析。

### 3. 定义

在详细描述本发明前，将定义在本发明的上下文中使用的一些术语。除了这些术语，根据需要，也在说明书的别处定义出现的其他术语。除非在本文中另有清楚的定义，在本说明书中使用的技术术语具有它们所述领域公认的含义。

“医学成像系统”是指任何能用于收集、处理和生成患者身体的一些或所有内部区域的图像的系统。典型地，此类系统包括生成和收集数据的装置，以及配置为处理和分析数据并且经常生成表示数据的输出图像的计算机。用于生成和收集数据的装置包括那些非侵入性的装置，如，磁共振成像 ("MRI") 仪、正电子发射断层显象 ("PET") 仪、计算机化轴向断层显象 ("CAT") 仪、超声仪等，以及非侵入性地生成和收集数据的装置，如，内窥镜（用于从机体的腔或管腔内传送可视图像）和有检测能力的导管。随后从这些装置中收集的数据被传送到处理器，该处理器至少在一些情况下能够用来产生患者机体的一个或多个内部区域的图像。随后受过解释图像训练的卫生保健专业人员检查和解释图像以形成诊断或预后。

根据本发明，“可以取得专利的”组合物、方法、设备、制品、或改进意味着在进行分析时主题满足于所有法规关于可专利性的要求。例如，关于新颖性、非显而易见性等，如果之后的研究表明一个或多个权力要求包括一个或多个不具有新颖性、非显而易见性等的具体实施方式，那么该权利要求，受限于“可以取得专利的”具体实施方式的定义，可以特定地排除不可以取得专利的具体实施方式。同时，本文所附的权利要求将被解释为既提供最广的合理范围也保持它们的有效性。而且，从本申请被提交的时间或者作为专利出版的时间到一个或多个所附权利要求的有效性有疑问的时间，如果一个或多个关于可专利性的法规要求被修改，或如果用于评价关于可专利性的特定法规要求是否得到满足的标准发生变化，那么将以下述方式解释权利要求：（1）保持它们的有效性以及（2）在这些情况下，提供最广泛的合理解释。

术语“治疗”意思是指疾病或病症的任何治疗，包括：预防或防治疾病或病症（也就是说，导致临床症状（或可以产生或有助于症状的潜在过程）不发展）；抑制疾病或病症（即，阻止或抑制临床症状的发展，或抑制一个或多个有助于可以产生症状的病变的潜在过程的发展）；和/或减轻疾病或病症（即，引起临床症状的消退；或引起一个或多个有助于症状的过程的消退）。应当明了由于最后的诱导事件或多个事件可以是未知或潜在的，所以并不总是可以分辨“预防”和“抑制”疾病或病症。因此，术语“预防”可以理解为形成包括“预防”和/或“抑制”的“治疗”类型”。术语“防护”因此包括“预防”。

## 发明内容

本发明的目的是提供用于自动化检测和分析（如果需要）患者脉管系统的一个或多个区域中的斑块的方法、软件和系统，其中从

医学成像系统获得数据,或者通过起始检测或数据收集过程(如(但不限于)那些能够用来生成图像的过程)来获得数据。

因此,一方面,本发明涉及评价患者脉管系统的动脉粥样硬化程度的自动化方法,该动脉粥样硬化存在于至少部份的患者脉管系统中,更常见的是存在于整个一根或多根血管的部分中,尤其是那些提供血液到器官如脑、心、肾、肝、肺、肠、膀胱、胃、卵巢和睾丸的血管,以及供应血液到外周如胳膊和腿的血管。用于分析的优选血管包括颈动脉、冠状动脉、以及主动脉。虽然本发明的方法可以用来检测和分析各种动物体内的脉管斑块,但这些方法最常应用于人。

通常,本发明的方法包括计算处理可处理的数据,该数据来源于患者脉管系统的至少一根血管的至少一个横断面(或其部分)并来源于医学成像系统,以确定是否血管(或至少在分析中的部分)含有至少一个斑块成分或与斑块的存在相关的组织。因此此类方法的性能允许评价一个或多个与患者脉管系统的至少一部分中的动脉粥样硬化相关的测量。

在优选的具体实施方式中,这些方法允许测定是否血管含有斑块,尤其是易破裂的斑块。对于特定的横断面,分析的数据可以包括部分或所有的原始收集的数据。用来获得原始数据的医学成像系统可以是侵入性或非侵入性的成像系统。优选的非侵入性成像系统包括一个和多个MRI、CT、PET、温度记录法或超声设备。也可以使用包括多个非侵入性成像功能的设备。优选的侵入性设备包括装备有一个或多个传感器的导管。实例包括用于静脉内超声、血管镜检查、脉管内MR和温度记录的导管。也可以将用侵入性和非侵入性成像技术获得的数据合并,以供分析。同样地,也可以包括其它或另外的数据,例如,利用对比剂所得的数据,其标记对于一个或

多个组织、细胞类型、或配体特异的部分，其，例如，包括健康或患病的脉管系统的组织或成分，包括斑块或其组分。

基于 MRI 的方法代表优选的一组具体实施方式。在这样的具体实施方式中，MRI 设备用来产生原始的磁共振数据，从这些数据可以得到可处理的磁共振数据。通过一个或多个不同的射频脉冲序列来实现的一个或多个不同的成像模态能允许基于随后的分析来鉴别不同的组织和组织成分。通过这样的模态产生的优选数据类型包括 T1-加权数据、T2-加权数据、PDW-加权数据和 TOF-加权数据。也可以结合通过组合一个或多个这些数据类型和其他数据类型所生成的数据。

在实施本发明的方法时，可能希望对数据进行预处理和/或规格化。在任何情况下，对可处理数据进行计算处理以确定在横断面（或其部分）的区域中的血管是否包括动脉和斑块组织或其组分。在优选的具体实施方式中，组织或成分类型确定是通过计算机将在数据中识别的不同组织类型和一个或多个统计分类装置进行比较来完成。这样的分类装置可以通过使用任何合适的方法并利用已知结果的数据（如：通过手术后的组织检查、直接组织观察、或一个或多个专家的标记）来开发，其中合适的方法包括逻辑回归、决策树、非参数回归、Fisher 判别分析、Bayesian 网络建模、以及模糊逻辑系统。优选进行筛查的成分和组织包括肌肉、外膜、钙沉积物、胆固醇沉积物、脂质、纤维斑块、胶原、以及血栓。

在优选的具体实施方式中，特别是那些其中使用从多重成像模态或成像设备获得的数据的具体实施方式中，数据被转换为普通格式。它也优选通过计算被引导配准（registration），通常使用标志，假如代表一种物理特征（例如，动脉分枝点如颈动脉杈）或计算特征，如从被处理的数据计算得到的管腔形心（centroid）。在一些具体实施方式中，可以计算处理至少在某个区域的一部分的血管的三

维立体模型，而所述区域是由待分析的空间距离最远的横截面所限制。也可以进行许多其他分析或操作，包括计算总体斑块体积或含量、以及斑块的位置和/或组成等。取决于进行的分析或操作，可以把分析的结果输入一个或多个输出文件和/或传送或转移至系统不同的位置以便存储。可替换地，数据可以传送至不同的位置。

本发明的又一个方面涉及评价治疗方案的有效性或确定治疗方案。这样的方法利用了本发明的斑块检测和分析方面，联同提出或确定医疗方案，视情况而定，其依赖于斑块检测、以及优选地分类、分析的结果。在一些具体实施方式中，医疗方案包括给予药物，该药物预期可以随时间稳定或减少患者的斑块含量。如果需要，可以通过跟踪分析来评价医疗方案的效果，优选通过进行额外的斑块检测，以及优选根据本发明的分类、分析。应当明了本方法不仅可用于提出认可的治疗策略，而且可用于开发新的策略。例如，这些方法能用于评价研究性治疗的临床效力，包括那些与被评估的用于治疗心血管和/或脑血管疾病的药物相关的研究性治疗。

本发明的另一个方面涉及计算机程序产品，其包括计算机可用介质（计算机可读程序代码体现在其中），其中计算机可读程序代码配置为在适合于执行该计算机可读程序代码的计算机上实现根据本发明的自动化方法。

配置为执行这种计算机可读程序代码的计算系统是本发明的另外的方面，如用于执行此类方法的商业模型，例如，ASP 和 API 商业模型。例如，在 ASP 模型中，医学成像系统和配置为执行本发明的计算机可读程序代码的计算机系统存在于不同的位置。通常，计算机系统存在于物理上远离许多成像中心的每一个的计算中心，每一个成像中心包括能够产生原始数据的医学成像系统，而从原始数据可衍生可处理数据。在优选的具体实施方式中，至少成像中心之一通过远程通讯线路传送原始数据到计算中心。

关于计算机系统，它们通常包括适合于执行本发明的计算机可读程序代码的计算机、和计算机连接的数据存储系统、以及可选地可操作地连接到计算机的通信接口，该接口用于接收计算机要处理的数据，或用于在计算机处理以后发送数据。

## 附图说明

通过以下详细描述和代表本发明的某些优选具体实施方式的附图，本发明的这些和其它方面以及具体实施方式将变得显而易见，其中附图可以总结如下：

图 1 是一流程图，其展示本发明的若干优选具体实施方式的概要。

图 2 有两个图片，A 和 B。图片 A 是从原始磁共振数据(DICOM 格式)生成的图像，原始数据获自商用 MRI 设备，该图片显示自表面线圈的照度梯度(illumination gradient)。图片 B 表示如展示在图片 A 中的在经直方图均衡化后的相同图象。

图 3 有四个图片，A-D。图片 A 表示从使用 T1-加权(T1W)模态获得的数据衍生的 MRI 图像。图片 B 表示从使用 T2-加权(T2W)模态获得的数据衍生的 MRI 图像。图片 C 表示从使用 PD-加权(PDW)模态获得的数据衍生的 MRI 图像。图片 D 表示体内 T1W、T2W 和 PDW 图像的多模态配准结果。

表 4 是一流程图，其表示在本发明的范围内有用的预测模型的过程。

表 5 有四个图片，A-D，其说明从 MRI 图像进行数据标记的过程。如将明了的，包括 MRI 图像的图像数据可以通过使用不同的规程(模态)收集的数据来生成。在本附图中，图片 A 表示人动脉横



断面的 MRI 图像，其是利用三种标准的 MRI 成像模态加以成像：质子密度加权 (PDW)、T1 松弛时间 (T1) 加权 (T1W)、以及 T2 松弛时间 (T2) 加权 (T2W)。为了容易目视解译，可以组合这些 PDW、T1W 和 T2W 图像 (分别是 **510**、**520** 和 **530**) 以生成伪彩色的组合 MR 图像 **540** (绿色=PDW, 红色=T1, 蓝色=T2)，其展示在图片 B 中。在显示在图片 B 中的组合图像中，多重对比规格化的灰度图像 **510**、**520** 和 **530** 分别被线性描绘成绿色、红色和蓝色信道，其中在每一个颜色信道中用黑色描绘零点和用白色描绘 **255** 以产生有色的组合图像以及用 MATLAB 进行三维处理。具有相似化学和环境特性的组织趋于有相同的颜色。关于组织类型的额外标记包括解剖位置 (例如，肌壁的内侧或外侧，即血管的内侧或外侧) 和结构 (texture) (例如，肌肉倾向于为纹状，而软的脉管斑块通常显示为“斑状”)。放射学专家经常能够通过对此类数据进行详细的人工检查来区分出纤维化的和易损的脉管斑块，但这样的工作极端费时并且主观。为了开发用于分类脉管斑块的自动化系统，模型必须用已知的实例 (“基础真值”) 来 “培训”。可以培训模型来模拟专家的操作，但优选用尽可能客观的标准，如使用组织的病理切片进行确证，来标记这些组织、或用来产生图像的数据。图片 C 显示动脉横断面的组织病理学 (基础真值)，该动脉横断面用来生成示于图片 A 和图片 B 中的图像。图 5 的图片 D 表示用于模型培训的标记图像，其中用不同的目标颜色来对每一个感兴趣的组织类型进行标记。动脉肌肉 (中膜，**565**) 是粉红色；外膜 (筋膜或胶质，**570**) 是亮黄色；血栓 (凝固血，**575**) 是红色；纤维状脉管斑块 (**580**) 是浅黄色；脂质 (**585**) 是白色，而管腔 (**590**) 是黑色。

图 6 有三个图片，A-C，并且代表数据标记的另一个实施例。图片 A 表示两个动脉横断面的伪彩色组合 MR 图像 (**610**)。MR 图像 **610** 是通过组合使用三种 MRI 模态 PDW、T1W 和 T2W 所产生的灰度 MR 图像而产生，如连同示于图 5 的伪彩色图像所描述的。

图片 B 描述的是动脉横断面的组织病理特性。图片 C 表示的是标记图像 (**630**)，类似于图 5 的图片 D 中的 MR 图像加以标记。

图 7 有三个图片，A-C，并且表示使用 K 均值 (K-means) 聚类算法处理的图像。

图 8 有两个图片，A 和 B，其说明当相对于标记基本真值 (每个图片的左部) 测量时优选具体实施方式的性能。

图 9 包含一张表 (表 A) 和三个图，其总结了用于检测脉管斑块、粥样斑中成分 (即，脂质) 和肌肉组织的三种预测模型的性能。表 A 表示基于最大 Kolmogorov-Smirnov 统计 (Max-KS) 和 ROC 曲线 Gini 系数测量结果的 RIPNet 模型的性能，其中 ROC 曲线显示在图中的其它地方。

图 10 有三个图片，A-C，表示对于低质图像本发明的优选具体实施方式的性能，其中低质图像由模型开发过程所提出。

图 11 有两个图片，A 和 B，其显示图片 A 中的图像内腔中心转换成图片 B 中的极坐标。这种转换用来改善总边界检测算法的性能。

图 12 有两个图片，A 和 B，其表示对于两个动脉横断面的组织分割算法操作的结果。一旦施行组织切割，被错误地标记为血管壁外部的脉管斑块成分的像素被消除，从而减少了假阳性。而且，脉管斑块载量的估计能通过比较被归为脉管斑块的像素和壁内像素数目的比率来获得。在这些实施例，脉管斑块载量估计分别为 28% 和 62%。

图 13 表示在颈动脉杈区域的部分颈动脉的三维图。在该模型中，示出了动脉壁 (**1920**) 的内边界和血管腔 (**1940**) 内的硬脉管

斑块（1930），而没有示出动脉的外边界。在动脉壁（没显示）的内表面（1920）和外表面之间的脂质（1910）表示为红色。在模型中硬脉管斑块是米色（colored beige）。

如本领域技术人员将明了的，在附图中表示的具体实施方式仅是说明性的而不是描述本发明的实际范围。

## 具体实施方式

在详细描述本发明之前，需要理解的是本发明没有限定于所表述的特定的成像技术、方法和系统，因为这些技术、方法和系统可以变化。也需要理解的是本文使用的术语仅是为了描述特定的具体实施方式，而不是限制本文所描述的发明的范围。

本发明涉及自动化、客观的方法和系统来测定和分析在患者脉管系统的一个或多个区域的斑块。一般来说，本发明的方法包括对比从一维、二维或三维图像（利用医学成像系统（或对于这类系统的数据收集先驱）所获得）衍生的数据，以相对于数据库对患者进行检查，其中数据库包含信息，所述信息使得能够对来自患者的数据进行分类并检测脉管斑块（如果存在的话）。进一步的比较使得能够分析脉管斑块，例如，如果需要可以分类（如：分类为稳定的或易受损的脉管斑块）。模式识别技术用来进行这些比较。单独这种信息或这种信息与患者其他数据结合可用于各种目的，例如，确定治疗过程，对患者发生其后不良事件的危险（如中风或心脏病发作）进行分层。可用于实施本发明的成像技术是那些可以用来生成三维血管图像的技术，并且包括 CAT、PET、MRI 和超声。目前，MRI 是优选的。

实际上，通过把患者送入 MRI（或其他成像）中心，该中心会把患者放入能生成进行随后分析所需要的基本输入数据的成像装

置中来获取患者的数据。为了实现本发明，在成像中心不需要附加的硬件。一旦收集到原始数据，在优选的具体实施方式中，数据被发送（如，以一个或多个加密电子数据文件的形式经互联网）至分析中心。然后，通过使用成组的一个或多个统计分类装置把患者的数据模式和数据库进行比较，该数据被自动处理以形成个体化的产品。然后可以制备个体化的患者产品并送给需要的医生。在优选的具体实施方式中，患者产品提供如患者心脏的脉管系统的三维视图，例如，三维视图可以表明总脉管斑块的位置和易于破裂的脉管斑块子集的位置。它还用来量化单个脉管斑块、总脉管斑块、单个易损脉管斑块、和总易损脉管斑块的体积。当在某时间段内用来为给定患者，尤其是接受动脉粥样硬化症治疗的患者，进行大量分析时，本发明的方法和系统可以用于评价治疗的效力。例如，是否治疗已减轻患者的总脉管斑块的含量（和/或减小脉管斑块含量的发展（或预计发展）的速率）；是否已减小易损脉管斑块的百分比或量；是否特定脉管斑块的组成随时间而变化（如变得更加稳定或较小稳定等）；等等。

本发明的方法可以容易地体现在软件、硬件或其组合中，以便以用户友好和可重现的方式对动脉粥样硬化性（AT）病变提供自动化、非侵入性和客观的检测和分析（如脉管斑块鉴定和分类）。本发明可以使研究人员、医生、以及患者从现存的疾病管理和/或治疗策略中容易地获取增加的益处。对于在发达国家中构成发病率和死亡率的单个主导原因的这类疾病，这些重要的诊断和预后方法和系统因此将改善治疗和结果。

### 1. 用于脉管斑块检测和分析的自动化方法

一般来说，本发明的方法是基于利用医学成像系统获得的患者数据的计算分析，以确定是否患者在他/她的至少部分的脉管系统中患有动脉粥样硬化。为了检测脉管斑块，利用统计分类装置，计算

机对数据进行处理和比较，以确定是否所分析的血管的一个或多个区域包括与脉管斑块的存在相关的（也就是，已知伴随的）至少一个组织。如果需要，脉管斑块（如果存在），也可以加以分类，例如分类为稳定的或易损的脉管斑块，这取决于在脉管斑块区域中鉴别的组织。而且，也可以对如脉管斑块的体积、脉管斑块的含量、疾病的发展、治疗的效力等进行评价。

首先，利用医学成像系统收集患者身体（尤其是所有或部分血管）的至少一点、一条线、一个平面、一个横断面或三（或更多）维图像的原始图像数据。如在本文中所使用的，“横断面”的含意是其中体现的实际数据可以指更少或更多量的数据。优选的医学成像系统是非侵入性系统，并且包括MRI设备。然后将从成像设备收集的原始数据转换成适合计算机分析的形式。利用计算机进行分析，即对至少一个、以及优选多个不同的健康和已知在脉管系统中存在疾病组织，对给定横断面的已处理数据和至少一个、以及优选若干个统计衍生的分类装置或预测模型进行比较。这样，能够拼装至少一个血管的至少一个横断面的模型。当获得一些或许多横断面的数据时，可以拼装更大的模型，其跨越由各种横断面限定的区域。如果需要，获得的模型可以用来重建所分析的血管区域的三维模型，这种模型可以描绘血管的各种特征。例如，该三维模型可以显示血管中脉管斑块的位置。这种模型也可以用于计算血管一个或多个区域的狭窄程度，以及在血管特定区域内的脉管斑块的体积。可以使用任何适宜的方法计算脉管斑块的体积。例如，可以计算不存在脉管斑块的血管腔的总体积，正如可以计算在有脉管斑块存在的情况下在所述区域的血管腔的体积。差值能用来表示在所述区域中的脉管斑块的估计体积，以及也能够容易计算狭窄的程度（例如百分比）。同样，可以确定脉管斑块的含量，正如可以确定疾病的其他临床度量。

## A. 典型系统配置

使用 MRI 分析作为典型的实施例，根据本发明的用于脉管斑块检测和分析的系统的优选具体实施方式的全部设计示意地说明在图 1 中。如将明了的，系统的各种部件优选为模块化的，以便于可以更新或修改一个或更多部件而无需更新或修改整个系统。而且，许多显示的步骤是可选的，但是已被包括在内以便描述本发明的方法和系统的当前优选的具体实施方式。在给定的应用中去掉一个或多个这些可选的元件、步骤、或过程是允许的。

如图 1 所示，处理从 MRI 中心或其他装置 (110) 收集的患者 MRI 数据开始。收集的原始数据 (105) 被传递到脉管斑块检测和分析系统，该检测和分析系统可以作为体系的部分存在于收集数据的装置中或者存在于不同的装置中。为了在另外的装置中进行快速数据处理，数据优选用电子传送，例如，以加密数据文件的形式经互联网传送至包含一个或多个计算机的装置，其配置成处理数据，以便检测和分析（如果需要的话）脉管斑块。优选地，测试原始图像数据以确保其符合最小的质量标准（数据质量分析 120），例如，通过计算种群稳定性指数。如果数据没有充分的质量（以及在本发明的具体实施中不能提供充分的质量）来提供可靠的输出，则不能进一步处理数据，并且信息优选传送到成像中心以告知对供分析的原始图像数据的拒绝。如果需要，初始原始数据的另一个副本可以被再传送，或，可替换地，可以收集另一原始数据集 (105) 并再传送以供分析。

满足质量确保参数后，可以允许原始数据作进一步的处理。在优选的具体实施方式中，原始数据被预处理和/或规格化（步骤 130），然后进行计算分析以初步鉴定血管的总体结构 (140)。当两个或更多数据的模式可用于分析时，随后利用任何合适的配置为基于计算机的实施的算法将截面和不同的数据模式引导配准 (150)。然

后可以进行图像变换、结构处理、以及变量计算（即，图像处理，**160**），其后可以利用统计分类装置或预测模型对数据进行分类以确定组织分类（**170**）。然后血管中总体结构边界可以测定（步骤**180**），并且从各种数据拼装血管的三维重建（**185**）。此后，进行病变（这里指的是脉管斑块）诊断，其后可以产生血管的三维模型，如果需要的话，随同诊断/预后报告和/或标记图像（**195**）。如果需要的话，随后结果被送至指定的接收者，例如，医生、诊所、或供其后检索的数据存储系统。

上文描述并示于图 1 的系统的一些步骤会在下文更详细地加以描述。

#### i. 数据输入

通过任何合适的方法，可以将患者的原始图像数据（**105**）提供给根据本发明的脉管斑块的评价系统。一个这样优选方法是 ASP（应用服务供应者）模型，其中患者的原始图像数据（**105**）经由安全的互联网连接从成像装置传送。另一个模型是（应用程序接口，或 API）模型，其中脉管斑块的评价系统嵌入就地安装在成像装置中的软件包中。

#### ii. 图像处理和格式化

在优选的具体实施方式中，需对原始图像数据进行质量保证检查以确保数据满足最小的数据质量标准。然后预处理满足这些标准的数据。例如，由于使用不同的 MRI 设备，不同版本的设备控制软件等，使得从不同的 MRI 成像装置获得的数据可以是不同格式的。虽然依照本发明，其他的格式也可以被采纳使用，但是用于 MRI 来源的数据的优选普通格式是 DICOM。同时，因为在各种 MRI 设

备和 RF 线圈之间硬件有差别，如果需要，可以以产生相对无干扰噪声和无失真数据的方式来补偿分辨力和比例。

已知 MRI 信号强度随着与表面 RF 线圈的距离而下降 ( $1/R^2$ )。结果,如图 2 所示,从原始 MRI 数据产生的图像表现出“照度梯度”。如果需要,可以使用可用于校正此效果的任何现在已知或将来开发的算法。合适的方法包括直方图均衡化(Gonzalas and Woods, *Digital Image Processing*, 1992 Addison Wesley), 以及使用小波来模拟 RF 线圈功能。可以获得能用于校正此效果的其他算法,其是作为商用图像处理软件工具的一部分,如 MATLAB (Mathworks,Inc.,Natick,MA)。也参见 Han, *et al.* (2001), *J. Mag. Res. Imaging*, vol. 13: 428-436。

加强的图像分辨很少依靠绝对的(而不是相对的)像素强度,主要因为强度经常取决于特定的条件和用来收集数据的成像设备。因此,通常有必要在每个相应的图像中把数据规格化为它们的最高像素强度。典型地,来自于每一个数据收集模态(如, T1, T2, PDW, TOF 等)的数据被独立地规格化,以便每一个模态的数据有相同的动态范围。然而,比较绝对模态内强度差值的附加变量,可以使用非规格化的数据来生成。多重模态变量,如 T1 和 T2 的比例,例如,可测量规格化量的比例。

在一些情况下已发现,像素强度的动态范围会缩小,例如一些体内颈动脉图像。到序列的端部(最接近头部),观察到的分辨率可能相当差,这可能是使用局部脖圈(neck coil)的结果。然而,根据应用,例如,与脉管斑块分类或微观结构(如血管再生或纤维盖厚度)鉴定相反,对于估计整个脉管斑块含量,低分辨率图像可能仍是有用的。而且,血抑制脉冲序列也能提高分辨率(Yang, *et al.* (2003), *International J. of Cardiovascular Imaging*, vol. 19: 419-428),



正如使用几个模态收集数据一样。例如，脉管斑块的纤维盖能在 TOF 图像中被很好地分辨出来。

### iii. 初步总体结构鉴定

为了检测和分析患者体内成像横断面上的脉管斑块，经常需要鉴别试图要分析的血管。总体组织鉴别允许从 MRI 切片中提取研究区，如血管，以供分析。例如，使用形态学技术来鉴别脉管的管腔可以容易完成这一工作。当然，也可以单独或结合腔检查使用其他总体形态特征（如，动脉肌肉、外膜等）的鉴别。当使用腔检测时，一旦起始腔位置被确定，随后的图像切片可以使用先前切片中腔位置的估计来进行腔位置的初步估计。一旦找到，对每一个切片，优选重复再估计腔中心（即，腔的形心）。在连续的切片中，尤其在有不规则特性的疾病组织的范围内，为了避免形心估计误差的混杂可以使用附加试探式算法，如在更远侧轴向位置再配准和在切片间的插入。

### iv. 图像配准

进行多模态 MRI 扫描要求的时间间隔可以产生模态间和模态内对准和配准错误，这是由于患者移动、心跳、呼吸、动脉扩张等所造成的。对于颈动脉图像，每个扫描平行获得许多图像切片，例如，优选 12 - 20 个切片，目前常规的商用 MRI 设备需要 3-4 分钟。对多模态图像需要附加扫描。因此，对于使用常规的 MRI 设备进行的颈动脉成像，全过程也许（目前）需要 3-20 分钟或更长时间。虽然控制（gating）心跳或呼吸循环并不产生更多有益于颈动脉成像的效果，但更长的扫描时间（如，正如使用多模态进行扫描可能需要的）可以增加在扫描过程中患者移动的可能性。对于冠状动脉成像，因为搏动心脏的运动，控制可以以 EKG 为基础来收集原始的

磁共振数据，虽然这样做经常明显地减慢每个模态的过程，每个模态大约 10 分钟时间是常见的。

当使用多重模态时，最可能要求模态间配准或对准。利用腔形心的直接对准可以得到合理的配准。然而，由于在所有 MRI 模态中血液的高对比，形成“腔检测器”来把图像集中在重要的参照点或标记上是非常不必要的。对腔进行检测会允许确定总体腔边界的位置，然后总体腔边界能用作图像配准的起始参考点。Woods, *et al.* (1998), *Journal of Computer Assisted Tomography*, vol. 22: 139-152。虽然通过刚性体平移和旋转（见图 3）可以获得令人满意的对准，但也可以使用考虑到组织变形（例如，起因于血压变化）的其它更复杂的方法。参见，例如，Dhawan, A. (2003), *Medical Image Analysis*, IEEE Press Series in Biomedical Engineering。

如将明了的，涉及更精细对准（例如，像素对准）的方法，优选使用量度，借此可以量化配准的质量。这种量度可以同规格化的交叉关联同样简单，或者可以更复杂，例如交互信息的最大化。Viola and Wells (1995), *Alignment by Maximization of Mutual Information*, International Conference on Computer Vision; Wells, *et al.* (1996), *Med Image Anal.*, vol. 1 (1): 35-51。重要的是要注意到，当对准图像或数据集（使用不同的成像模态所得到）时，参考图像和被对准的图像经常表现出不同的特性。因此，如果使用相同模态收集两个图像或数据集，那么对准最大化标准不能呈现如期望那样的清楚的峰。

对于垂直配准，用来分析的脉管管腔优选使用来自不同模态的对准切片，其接近常见的解剖参考点、计算上的拟制（如，管形心），或其他标记。然后随后的切片从这个常见点可以容易地被对准。例如，在颈动脉成像中方便的参考点是颈动脉杈。实际上，在以下实

施例中描述的分析使用颈动脉权作为轴向参考点。在不同模态中切片间间隔也可以要求线性插值算法。

#### v. 图像处理

在优选的具体实施方式中，可处理的数据（即，成一定配置以供计算机操作的数据）经过图像处理算法，以便去除噪声以及合成结构特征和其他感兴趣的变量。虽然非参数回归模型（如神经网络或径向基础函数）可以用来评估任何随机的，非线性判别函数。Cybenko, G. (1989), *Mathematical Contl. Signal & Systems*, vol. 2: 303-314 ; Hornik, et al. (1989), *Neural Networks*, vol. 2: 359-366; Jang and Sun (1993), *IEEE Trans. Neural Networks*, vol. 4: 156-159。实际上，将任何已知关联合并到变量集内是有用的，这可以简化优化问题。普通技术包括变量线性化和变换或组合变量来寻找非线性关系等。例如，在建立模型来识别地震信号时，首先将时间系列变换成频域是非常更有效的。Dowla, et al. (1990), *Bull. Seismo. Soc. Amer.*, vol. 80 (5): 1346-1373。然后，图像处理的全部目标是建立输入图像的转换。所采用的图像处理操作的类型大概分成若干（不是互相排斥）类，此分类基于它们的数学目的：噪声缩减；维数缩减；结构或特性检测；以及衍生变量（通常利用专家专业知识设定，虽然它们能使用数学/统计学技术来定义）。在以下的实施例中描述了变量和转换的实例，其证明可以增强根据本发明的脉管斑块分类系统的性能；然而，在本技术领域已知的其他图像处理技术也可以适用于实施本发明。

#### vi. 组织分类

图像处理（160）后，经变换的数据被导入统计分类装置中以将图像中每一个像素分为属于若干组织之一，包括脉管斑块的成分。标记图像是对图像中每一个像素进行数学函数的直接处理。开

发预测模型的一种方法在下文实施例 2 中加以描述。然后在实施例 3 中提供了建立用于来自 MRI 图像的脉管斑块分类的预测模型的详细实施例。

### vii. 组织分割

图像分割是基于组织分类装置的输出进行，以便突出显示感兴趣的组织、狭窄程度等，以及抑制非相关特征。在许多情况中，区别脉管斑块成分和非病理性组织在解剖范围之外是不可能的。例如，硬脉管斑块本质上是疤痕组织，其主要组成是胶原，就像动脉筋膜。动脉壁外部的胶原是结构性的，并且肯定不是病理性的。同样，在检测和分析血管内斑块的范围内，脉管外部的脂质或钙沉积物并不具有临床意义。任何适宜的方法都可以用于此过程。在一优选的具体实施方式中，可以开发领域知识，正如某些变量作为离腔边界的径向距离的函数失去灵敏性。在另一个优选具体实施方式中，使用两级方法可以得到极好的结果，借此组织类型预测被通过第二、总结构处理模块。实质上，预测模型的输出被传入图像处理算法（例如，梯度流动和主动轮廓控制(Han, *et al.* (2003), *IEEE Trans. Biomed. Eng.*, vol. 50 (6): 705-710)来限定动脉肌肉的边界。然后边界外的所有像素可以不考虑作为脉管斑块成分或动脉壁（界定血管的内部）内的其它组织。使用“主动轮廓”算法或“横向震荡 (snakes)”算法的方法（Xu, P. (1997), *Gradient Vector Flow: A New External Force for Snakes*, IEEE Conference on Computer Visual Pattern Recognition; Xu, P. (1997), *Snakes Shapes and Gradient Vector Flow*, IEEE Transactions on Image Processing）是基于示于图 12 的体外数据加以说明。

本领域技术人员已知的其他分割算法也可以适用于在本发明范围内的应用。例如，使用基于规则的方法可以完成组织分割。示于图 11 和图 13 中的结果是使用这样的方法获得的。这类方法也能

连同边界检测方法一起使用，其中边界检测方法涉及寻找最小成本路径(Bishop, C. (1995), *Neural Networks and Statistical Pattern Recognition*, Oxford University Press)。在用来产生示于图 11 的结果的过程中，基于规则的方法用来将脉管图像（集中于腔）转换为径向坐标系统，其使固有径向的特性线性化。

#### viii. 三维重建

一旦标记 MRI 扫描的所有切片，如果需要，就可以产生动脉和脉管斑块的完整的三维模型。从商用 MRI 设备（例如：由 General Electric 制造的 MRI 设备）获得 DICOM 格式数据，直接从该数据检测总体结构（例如：腔和动脉壁外部）的算法可以用于这一目的。实施例 4 描述了此种模型如何生成的代表性的实施例。

#### ix. 病变诊断

病变诊断包括总体狭窄的尺寸和程度、脂质含量、脉管斑块的大小和体积、血栓和钙化等，病变诊断可以从血管的三维结构重建（体素（Voxels））来估计。当然，选择性地检测脉管斑块成分（如脂质）的成像模态可以用来从更少的数据生成有用的模型，因为可以要求更少的成像模态（如：T1, T2, PDW, TOF 等，在 MRI 分析的范围内）来生成检测和分析（如按易于破裂来分类等）脉管斑块的模型。

#### x. 数据输出

可以以标准化和惯例格式的形式提供系统的输出，以便包含可能要求或需要的信息，从而评审生成的结果。在一些具体实施方式中，输出将包括原始数据、预测模型标记的数据、三维模型、以及诊断报告，如果指示，输出数据包括危险因子和推荐的疗法。优选地，得到的输出直接送到系统，尤其是基于 API 模型的系统。在

ASP 模型的范围内，执行分析的计算机系统将自动地或在接受适当指令下传送输出文件到特定的地址。这样的地址可以是主治医生、放射科医师、和/或专科医师、被检查的患者、和最初传送患者数据的医疗成像装置等的电子邮件帐户的地址。

#### xi. 一般化和规格化

如将明了的，本发明的方法的自动化特性将便于开发标准化数据分析程序、格式等。并且，目前基于人类专家对成像数据检查的许多主观特性（因此可变性）可以通过实施本发明的方法和系统来去除。

#### xii. 其他考虑事项

如本文所描述的，MRI 可以用于鉴定脉管斑块的形态特征，例如脉管斑块的大小和纤维盖的厚度，并具有高度的灵敏度和特异性。而且，MRI 可以区别在所有主要动脉（颈动脉、冠状动脉、和主动脉）中易损和稳定斑块的斑块成分（如纤维盖、钙化、脂质含量、出血等）特性。已经使得成像协议得到改善以使运动伪像最小化。Worthley, *et al.* (2001), *Int'l J Cardiovascular Imaging*, vol. 17: 195-201; Kerwin, *et al.* (2002), *Magnetic Res. In Med.*, vol. 47: 1211-1217.

MRI 的优点在于能通过使用一些不同的模态来对结构进行成像。相同解剖组织的 T1-, T2-, PD-, 和 TOF-加权图像（分别是 T1W, T2W, PDW, 和 TOFW）可以是完全不同的，这取决于组织的化学成分和结构。例如，钙化、纤维组织、和脉管斑块内出血可以用 T2-加权图像来区分。钙化在质子密度加权（PDW）图像中是非常高密度的，而平滑肌能通过相对短 T 来很好地表征。飞行时间（TOF）加权图像能很好地区别脉管斑块内出血和富脂质坏死核心。造影剂可用来提高新脉管系统的检测，其是脉管斑块易损性的

另一个指示剂。而且，其他试剂，如标记抗体，包含对脉管斑块成分特异的靶向部分的小泡，也能用来提高或加入从根据本发明的用于分析的医学成像系统收集的数据。

目前，发明者已经确定，根据基于 MRI 成像的本发明的自动化方法的脉管斑块的检测和分析优选使用衍生自二个、三个、或四个不同成像方式（例如，T1，T2，PDW，和 TOF）或其衍生物（例如，T1/T2 比率）的数据，以便将脉管斑块成分和血管的其它组织区分开，虽然单个和其它多模态分析也在发明的范围之内。从多重对比获得的信息集成会有助于对脉管斑块的存在、位置、和组成作出更快速、准确、和可重现的评价。然后此种分析可以用来减少为测量和分类脉管斑块所需模态的数量以及可能导致具有更高辨别能力的 RF 序列的设计。同样地，使用对脉管斑块的特定成分特异的数据收集模式将减少初始数据收集的时间，正如成像设备的硬件、操作软件等的改善将减少初始数据收集的时间一样。

## 2. 应用

在破裂/侵蚀的人动脉粥样硬化损伤处的急性血栓形成对急性冠状动脉综合症的发作和动脉粥样硬化的发展起着关键的作用。已经清楚地确定了病理学证据：不是狭窄严重程度而正是脉管斑块的组成调节脉管斑块的易损性和血栓形成。如将明了的，本发明的方法和系统可以用于基于模式识别的自动化成像分析，其用于检测、测量、和分类体内动脉粥样硬化性斑块，以及总脉管斑块含量和相关的测量。在优选的具体实施方式中，使用 MRI 产生三维图像。自动化可以提供快速、客观的（不依赖于观察者）数据分析。这样的方法会有各种应用，包括检测和分析（如果需要）脉管斑块。分析可以包括，例如，量化脉管斑块的体积、确定脉管斑块的位置，和/或评估脉管斑块的组成。而且，脉管斑块的分析可以集中在脉管系统内的一个或多个区域，和/或通向有或没有已知心血管疾病（该信

息能帮助指导治疗，包括手术介入和药物治疗)的患者体内一个或多个区域或器官(如，脑、心脏、肾等)，评估总脉管斑块的含量(例如，在患者筛查、疾病治疗等范围中)，以及危险估计和分层。这些方法也可以用作标准的、客观的诊断和预后测量，因此便于在整个纵向研究等过程中对试验室之间的结果进行比较，从而在药物和其它治疗的临床试验中评估替代物终点(surrogate end point)，以及超越不同的成像设备。在临床配置中，这些方法也会极大地缩减诊断成本，该成本涉及测定狭窄的程度和检测易于形成血栓的脉管斑块，这些方法也会减少那些否则可能不得不使用更加侵入性诊断方法的患者的危险和负担，而同时提供比使用现有方法可得到的更多的有用信息。

#### A. 心血管疾病

因此，本发明可应用的一个方面涉及心血管疾病。已知在男性和女性中，心血管疾病是引起死亡的单个主要原因。在发达国家大约一半人群死于心血管疾病，而更多的人群会有和心血管疾病相关的并发症并伴随着较低的生活质量。单独在美国，每年花费在检查心脏和脉管斑块的产品上的花费就超过150亿美元。最近的发现表明，易损的、不稳定的脉管斑块破裂而引起心脏病发作和中风。值得注意的是，大约70%的破裂产生心脏病发作的脉管斑块来自于几乎没有脉管斑块的脉管系统区域。然而到此刻为止，还没有开发出客观、快速的方法来区别易损不稳定的脉管斑块和稳定的脉管斑块，其中易损不稳定的脉管斑块更可能破裂和引起可导致心脏病发作或中风的血栓形成。本发明通过提供非侵入性的、客观的、和快速的方法来检测和分析整个脉管系统中的斑块(尤其在脑、颈部、和心脏的脉管系统中的脉管斑块)从而致力于这类明显未能满足的需求。



### i. 术前病变诊断和患者筛查

当前美国心脏联合会的所有指导方针都是基于狭窄程度和症状状态，而没有涉及脉管斑块的组成。显然，更精确的术前诊断（例如，脉管斑块组成，如钙化、脂质含量、血栓形成、纤维盖的厚度等等）会明显地改善术前危险的评估，这使临床医生更可靠地评价比起药物介入来外科手术的相对危险。

### ii. 治疗

许多用于患者的心血管和脑血管预防措施和治疗是基于患者的心血管疾病（CVD）或脑血管危险的估计。为了达到本说明书的目的，将把 CVD 作为本发明一般涉及的动脉粥样硬化症的代表性实例来讨论。因此，Joint National Committee 的高血压指导方针，和 Adult Treatment Panel/National Cholesterol Education Panel 的胆固醇指导方针通过预期的 CVD 风险为治疗规定了适当性：也就是说，他们基于 CVD 危险评估，定义治疗阈值的百分比、或治疗开始时的血压或胆固醇水平。另外，他们通过预期的 CVD 危险，定义了治疗的目标：也就是说，治疗的进度、或血压或胆固醇的水平降到治疗应推进到的程度。

这在理论上已被证明，因为更高 CVD 危险的个体有更多危险要降低：在那些处于更高基线的 CVD 危险中，危险的相同分数的下降导致危险的更大绝对下降，同时更大的 CVD 危险降低提供更大的治疗成本效力（需要更少来治疗或预防 CVD 事件或死亡）；以及更可能的是（更大绝对的）治疗好处会多于治疗伤害。

当前 CVD 危险评估（基于其预测治疗阈值和目标）的方法没有包括关于易损的脉管斑块的信息。由于易损的脉管斑块是 CVD 危险的关键决定因素（可论证其是最重要的决定因素），以及由于

本发明使得能够以客观、自动化、以及容易使用的方式来检测和分析易损脉管斑块，所以大大提高了 CVD 危险预测（以及靶向不同端器官的危险预测）的准确性，从而允许显著改善治疗的命中（targeting）。

由于例如易损脉管斑块、总脉管斑块含量等的有效评价，改善的 CVD 危险预测（以及危险分层）可以对节约成本和挽救生命具有重要意义。对于那些真正处在危险中的患者改善的治疗命中将在相同花费下挽救生命，以及对于相同的挽救生命节约了花费。

脉管斑块的检测和表征也会允许更好地决定：谁值得药物治疗，以及什么样的药物治疗将最好地服务于特定的患者。这可以包括分配（昂贵的）的斯特汀类（statin）胆固醇降低药（例如，阿托伐他汀，辛伐他汀，普伐他汀，洛伐他汀，罗苏伐他汀，以及氟伐他汀），这些药目前占世界上任何处方药物治疗的最大花费，并具有每年 200 亿美元的市场，而且其使用随老化人群预期会显著上升。更一般地说，脉管斑块的检测和分析也可以为治疗方案改善治疗决定，这些治疗方案通过任何一组机制来影响心血管的危险，包括降低血压（如噻嗪类利尿剂，例如氢氯噻嗪， $\beta$ 阻断剂如阿替洛尔，血管紧张素转换酶抑制剂如福辛普利，血管紧张素受体阻断剂如厄贝沙坦，钙通道阻断剂如硝苯地平（硫氮<sup>萘</sup>酮和维拉帕米）， $\alpha$ 阻断剂如哌唑嗪（特拉唑嗪），和血管扩张剂如肼屈嗪），稳定脉管斑块（如通过一些斯特汀类药（statin）能够达到），降低脂质（如可以通过使用下述来完成：斯特汀类药；纤维酸衍生物如吉非罗齐或非诺贝特；烟酸或变体如烟酸控释制剂；胆汁酸螯合剂如考来替泊或考来烯胺；或胆固醇吸收阻滞剂如依泽替米贝），减少感染，和/或起抗血小板作用（如阿司匹林，氯吡格雷等）或抗血栓效果（如，组织纤溶酶原激活剂或链激酶）等等。当然，根据要治疗的特定患者和病情，可取的是联合一个或多个前述疗法，单独或结合其他治疗。

对那些最需要的患者，靶向外科治疗的改善可能是更重要的，因为与手术有关的潜在成本和危险应当仅由那些人承受，对于这些人来说，问题的真正危险超过手术的危险。对易损脉管斑块的评价可以大大提高是否患者真正有这种水平的危险确定。

### iii. 药物开发

在纵向研究中，研究人员已经使用非侵入性患者图像数据的人工评价来监测降低胆固醇药物的效力。然而，人工检查对于一般的临床诊断来说太昂贵了。相反，本发明的自动化方法和系统能用来对下述快速地产生产统计上可靠的估计：脉管斑块组成（如钙化、脂质含量、血栓形成、纤维盖的厚度等）、总脉管斑块含量、易损脉管斑块含量、易损和稳定脉管斑块的比例、或脂质沉积物，以用作临床结果的替代品（例如，破裂、中风、MI），从而大大地减少研究的时间和成本。本发明提供的另一个主要优点是显著地减少为证明心血管和脑血管药物的效力所需的患者的数量和随访的时间长度，这些药物包括那些进行临床试验的药物。例如，假设鉴于和斯特汀类药的使用有关的显著临床益处，进行任何新的试验（包括安慰剂）可能是不道德的。因此，为了说明相对于当前使用的心血管药物的显著优点，试验可以需要至少几千患者和 3-5 年的随访。

### iv. 增强的诊断

本发明的方法和系统的另一个应用涉及向患者和医生提供优良的诊断和预后工具。在这方面，从应用本发明而生成的脉管斑块的检测和分析数据和结果可以和其他来源的数据结合来提供甚至更好的诊断产品和服务。例如，Framingham 心脏研究数据库已经广泛用来建立得分卡以估计心血管疾病（CVD）的危险。从长时间追踪 5,209 位受治疗者（从其已获得许多预测变量），已开发了这一意义重大的数据集，所述预测变量包括年龄、性别、胆固醇和高血

压的量度、人口统计因素、药物治疗、糖尿病的状态、酒精消耗、吸烟史、心脏病史（例如心肌梗塞（MI）、心绞痛、心动过速、和心动过缓）、血管再生、冠状动脉旁路移植手术、中风、血中分析物的水平（如，肌酐、蛋白质）、个性类别（例如，A型）、和许多“新出现的”危险因素，包括C-反应蛋白质水平、VCAM, ICAM黏附分子、和血中其他物质。访问的结果包括死亡（特殊原因），以及心血管事件，包括MI、中风、和猝死。

在 Framingham 数据集中没有 MR 扫描，然而，来自最近纵向斯特汀类药物试验的体内患者 MRI 包括 MRI 图像和有关的患者历史和危险因素（如血压、胆固醇水平等），其是可以得到的。例如，在 Mt. Sinai 的研究小组收集 MR 图像，间隔为 6 个月，时间超过两年（Woods, et al. (1998), *Journal of Computer Assisted Tomography*, vol. 22: 139-152)。通过将使用本发明的方法得到的结果和与 CVD 相关的一个或多个其他数据点结合在一起，可以实现改进的甚至更好的诊断程序。

## B. 中风

对动脉粥样硬化性病变的组成的更好了解也会便于更精确的患者中风的危险分层，这有利于选择合适的疗法。大约 25% 的中风和颈部颈内动脉的闭合性疾病相关。治疗方案包括抗血小板疗法、动脉内膜切除术、安装支架（stenting）、和血管成形术。这些治疗中，对于晚期颈动脉病灶，颈动脉内膜切除术（前摄的手术去除）是优选的治疗方案，在美国每年进行的这些手术超过 120,000。一些大的临床研究，包括北美有症状的颈动脉内膜切除术试验（NASCET），无症状的颈动脉粥样硬化症研究（ACAS），和欧洲颈动脉外科手术试验小组（ECST），已经表明在一定的限度内此手术可明显地减小中风的危险。对于 70% 狭窄的有症状的患者，在两年中中风的危险的总的降低估计为 17%。外科手术也增加

了手术期间事件的危险；在外科手术患者中死亡率从 0.3%增至 0.6%；严重的中风从 3.3%增至 5.5%；以及脑血管事件从 3.3%增至 5.5%。对于超过 60%狭窄的无症状的患者，中风和手术期间中风或死亡的合计危险估计达到手术患者的 5.1%，比较来说，对于药物治疗的患者达到 11%。而且，有证据显示，来自试验的相反的结果评估低估了在实际应用中不利结果的可能，这进一步增加了确定对哪些患者真正有益的重要性。

本发明的方法的使用将可以更好地对患者进行评估，以便可以采用合适的疗法。并且，如同心血管疾病一样，筛查可以使患者在疾病发展的较早期被诊断，能够及早介入治疗并随时间更大程度降低危险。

### 3. 基于计算机的实现

以上描述的本发明的各种技术、方法、和方面能部分或完全利用基于计算机的系统和方法来实现。另外，基于计算机的系统和方法可用来增强或提高以上所述的功能，提高可以执行功能的速度、以及提供另外的特性和方面作为文中其他部分描述的本发明的一部分或除文中其他部分描述的本发明的那些特性和方面之外。现介绍根据上述技术的各种基于计算机的系统、方法和实现。

以上描述的本发明的各种具体实施方式、方面、和特征可以利用硬件、软件、或其组合来实现，以及可以使用有一个或多个处理器的计算系统来实现。事实上，在一种具体实施方式中，使用基于处理器的能执行本文相关描述的功能的系统来执行这些部分（元件）。示例性的基于处理器的系统的包括一个或多个处理器。每一个处理器和通信总线相连。各种软件具体实施方式是利用此示例性的计算机系统来加以描述。在本说明书中本发明的具体实施方式、特点、以及功能并不依赖于特殊的计算机系统或处理器结构或特殊

的操作系统。事实上，鉴于本文的描述，在相关领域中的普通技术人员会非常清楚怎样使用其他计算机或处理器系统和/或结构来实施本发明。

一般来说，基于处理器的系统可以包括主存储器，优选随机存取存储器（RAM），也可以包括一个或多个其他辅助存储器，其包括磁盘驱动器、磁带驱动器、可移动存储驱动器（如可插入或可移动的存储装置和磁带驱动器、CD-ROM 驱动器、DVD 驱动器、软盘驱动器、光盘驱动器等）。在可替换的具体实施方式中，辅助存储器包括其他数据存储装置以便于计算机程序或其他指令可以被调入或以别的方式载入计算机系统。

本发明的计算机系统还可以包括通讯接口（优选和远程通信网络兼容）以便于允许软件和数据传送到、传送自计算机系统，或在计算机系统和一个或多个外部装置之间传送。通讯接口的实例包括调制解调器、网络接口（例如因特网卡）、通讯端口，PCMCIA 槽和卡等。经通讯接口传送的软件和数据会是信号的形式，这些信号可以是电子信号、电磁信号、光信号、或其他能够经由通讯接口接收的信号。通常这些信号经由通道提供给通信接口，其中该通道传送信号并可以使用无线介质、电线、电缆、光纤、或其他通信介质来实现。通道的一些实例包括电话线、便携式电话链路、RF 链路、网络接口、和其他通信通道。

在本文中，术语“计算机程序产品”等等一般指介质如可移动存储装置、在磁盘驱动器中可以安装的磁盘、和通道上的信号。这些计算机程序产品向计算机处理器提供软件或程序指令。计算机程序（也称作计算机控制逻辑）通常存储在主存储器和/或辅助存储器中。也可以经通信接口接收计算机程序。计算机程序，当被执行时，能够使计算机系统完成如本文描述的本发明的特征。特别是，计算

机程序，当被执行时，能够使处理器实现本发明的特征。因此，计算机程序相当于计算机系统的控制器。

在使用软件实施本发明的具体实施方式中，该软件可以存储在计算机程序产品中或经计算机程序产品传送并使用任何合适的装置或通信接口载入计算机系统。控制逻辑（软件），当用处理器执行时，促使处理器去执行如本文所描述的本发明的功能。在其他具体实施方式中，主要以硬件或硬软件结合的方式，使用例如硬件组成部件如 PAL，专用集成电路（ASIC），或其他硬件组成部件来实施本发明的方法。目的是执行本文所描述功能的硬件状态机的实施对于相关领域的技术人员来说将是显而易见的。

## 实施例

以下实施例描述本发明的某些方面并帮助本领域的技术人员实施本发明。这些实施例决不是以任何方式限制本发明的范围。

### 实施例 1

#### 图像处理

本实施例描述了在本发明范围内用于处理图像数据的若干特别优选的技术，包括噪声抑制、降维、以及结构（纹理）处理。

#### A. 噪声抑制

处理合成的多重对比图像以便减少噪声和引入平滑（smoothing）。在每一个情况下，首先图像经中值滤波处理以去除具有脉冲特性的噪声，然后用自适应 Wiener 滤波器使其平滑，其中 Wiener 滤波器调节周围“N”像素邻域中的统计资料。平均值和方差用在像素位置  $n_1$ ,  $n_2$  的强度“a”来估计：

$$\mu = (1/N^2) \sum a(n_1, n_2); \sigma^2 = (1/N^2) \sum a^2(n_1, n_2) - \mu^2$$

然后这些估计值用来给 Wiener 滤波器<sup>i</sup>指定参数“b”:

$$b(n_1, n_2) = \mu + \frac{(\sigma^2 + v^2)}{\sigma^2} (a(n_1 + n_2) - \mu^2)$$

用上述系数 b 在每一个图像强度平面上完成 2 维卷积。

## B. 降维

因为要处理的数据点和变量的数目，为了使噪声效应减到最低，优选减少数据集的维数以产生更少但更具统计意义的变量。聚类分析仅是用来减少由成像装置产生的原始数据的噪声和维数的众多方法之一，其目的是保留原始数据的最显著的特征。K-均值聚类分析是聚类算法的一个实例。在这样的算法中，形成“K”类，其成员存在于（特征空间）的离每类的估计形心最小距离的位置。此方法在聚类形心（cluster centroids）处作最初的估计，然后根据更新的类成员数再估计那些形心。

作为 K-均值聚类算法基础的步骤是:

- i. 选择许多具有原始形心的聚类“k”;
- ii. 通过指定每一个数据点至其最近的聚类形心来把数据点分配进 k 聚类;
- iii. 计算聚类分配矩阵; 以及
- iv. 估计每一个聚类的形心。



重复步骤 ii-iv 直到达到停止标准,通常当成员停止改变聚类成员资格 (membership) 时。参见 Bishop (1995), *Neural Networks and Statistical Pattern Recognition*, Oxford University Press。示范性的聚类分析结果显示在图 7 中。从图中明显可见一些聚类和特殊组织类型有高相关性。结构聚类产生了表观上非常令人满意的结果,但统计上不如预测模型好。因此 K-均值聚类种类被用作预测模型的输入。在本技术领域已知的其他工具也可以用来缩减维数,包括把典型人工神经网络 (ANN) 的理论上非线性曲线拟合能力和分级技术的稳定性 (Bates White, LLC software, RDMS<sup>TM</sup>) 组合起来的方法。结果,仅对那些本身已知有一些预测能力并也可以在模型中体现大量最有用的基本非线性效果的输入进行估计程序 (routine)。这些步骤可以使 ANN 培训阶段集中于参数中比用其它方法需要的更低维数和更少非线性的问题。还有其他技术包括主分量分析、独立分量分析 (Bell and Sejnowski (1995), *Neural Computation*, vol. (7) 6: 1129-1159), 和局部信息量度(规格)(Haralick, R. (1979), *Proc. IEEE*, vol. 67 (5))。

### C. 结构测量

组织和脉管斑块成分可通过其结构直观上加以区别。例如,肌肉和胶原组织通常是有纹状的,而坏死核心显示出斑状。关于结构没有形式数学定义,但这些特征有数学相关物,例如信息内容、空间频率等。一个通常使用分类区分 28 个结构量度。这里,使用了两类结构量度,其是关于局部强度变化量和空间频率的统计资料。使用的统计像素量度是标准统计量,应用于各种尺寸邻域。离散余弦变换用来产生“x”和“y”取向的空间光谱能的估计。例如,细节丰富的像素区域在更高空间频率中有更大比例的能量。2D DCT 表达式是:

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \begin{cases} 0 \leq p \leq M-1 \\ 0 \leq q \leq N-1 \end{cases}$$

$$\alpha_p = \begin{cases} 1/\sqrt{M}, p=0 \\ \sqrt{2/M}, 1 \leq p \leq M-1 \end{cases}, \alpha = \begin{cases} 1/\sqrt{N}, q=0 \\ \sqrt{2/N}, 1 \leq q \leq N-1 \end{cases}$$

#### D. 导出变量

导出变量是从表现出预测特征的基本实体中综合生成的。属于此类的两个变量包括原始变量的乘积和比例以及三个数据类型的其他组合。发现有强分辨力的三种类型的导出变量包括：(i) “*胖型检测器 (Fat Detector)*”，定义为 T1/T2 比例，其在脂质检测中是有用的；(ii) *轴旋转: YCbCr*: 在这种格式中，亮度信息存储为单个分量 (Y)，而色度信息存储为两个色差分量 (Cb, Cr)。Cb 表示在蓝色分量和参考值之间的差别。Cr 表示在红色分量和参考值之间的差别；和 (iii) 局部环境变量，其基于离腔边界的几何距离。其他潜在有价值的图像特性的实例列在下面的表 1 中。

表 1 图像处理变量和变换的实例

分类	变换
结构量度	平均值、方差、偏量 (Skew)、中位数、内四分位数范围、最小值、最大值、标准偏差、范围、峰度、局部信息内容
结构量度	离散余弦变换、子波变换、取向
导出变量	比例 (如 T1/T2)，离开腔的距离、离开动脉壁的距离
导出变量	RGB 到 YCbCr 颜色轴旋转
降维	K-均值聚类分析、主分量分析、独立分量分析

#### 实施例 2

#### 预测模型开发

“模型”是数据或系统的数学或统计表示，被用来解释或预测新情况下的行为。模型可以是机械的 (通常使用在物理科学和工程中) 或经验/统计的 (其中模型预测的目的不是解释基本的因果关

系)。统计建模的两个相关应用是开发统计分类装置和预测模型。统计分类装置是在一系列的观测的基础上用来区分对象的种类。预测模型的目的是从当前观测或系列观测中预测结果或预测未来值。本发明采用两类模型：统计分类装置用来分类组织和斑块成分；而预测模型用来预测和例如心血管病（CVD）相关的危险。

模型开发的过程依赖于特定的应用，但在图 4 中示意说明的一些基本过程对于典型的模型开发工作是共同的。首先，必须建立建模数据集，包括一系列观测值（“模式”）和已知结果，数值，或与对应于每一观测（被称为“标记”或“目标”值）的类别。在图 4 中，这称为数据集构造 410。这个建模数据集用来建立（或“培训”）预测模型。该预测模型接着用来分类新（或未标记）模式。如下所述，模型开发经常是变量创造，选择，模型培养，和评价的迭代过程。

#### A. 数据集建立

建模过程中的第一步一般是把所有可获得的事实，测量值，或其它可能和目前问题相关的观测值集合到一个数据集组中。数据集中的每一记录与给定事件的所有可获得的信息相对应。为了建立一个预测模型，应当为数据集中至少一些记录建立“目标值”。数学上讲，目标值定义因变量。在 CVD 危险预测的应用范例中，目标可以使用从纵向临床研究中观测到的临床结果数据加以设置。在脉管斑块探测和分析（例如，分类）的范围内，目标对应于，例如，由人类专家标记或由组织检查确证的图像。图 5 说明了在一个这样的应用中所使用的的数据标记过程。在此例中，每对模式/目标通常称作范例，或培训例，其用来培训，测试、或确证模型。如将明了的，什么构成模式范例依赖于建模目的。

### i. 数据分离

如图 4 所示，模型的实施通常包括数据分离（步骤 **420**）。大多数模型开发工作需要至少两个，以及优选至少三个数据划分：开发数据集（用来建立/培训模型的数据）**427**，测试数据集（用来评估和选择单个变量、初级模型等等的的数据）**425**，和确证数据集（评估最终性能的数据）**429**。为了服务于此目的，初始数据被随机分为三个数据集，数据集不必有相等的大小。例如，数据可分成 50% 是开发（**427**），25% 是测试（**425**），和 25% 是确证（**429**）。起初利用开发数据（**427**）来开发模型。根据测试数据（**425**）获得的性能用来监控一些问题如任何过度拟合问题，即，模型应当对开发数据（**427**）和测试数据（**425**）都展示出可比性能。如果相对于测试数据（**425**），模型对于开发数据（**427**）具有较高性能，就调整模型直到模型达到稳定的性能。

为了确认模型将如所预期的那样在任何独立数据集上运行，理想地留出数据的一部分仅用于最终模型确证。确证（或“支持”）数据集 **429** 包括一系列不用来培训模型的范例模式。然后，可用完成的模型来评价这些未知的模式，以评估（评分）模型在评价新模式中如何运行。

此外，一些应用可需要另外的，“过时的”确证集，来确认随时间流逝模型性能的稳定性。对于如神经网络或遗传算法这样的更复杂的建模方法，附加的“数据分离”通常是必需的。例如，一些建模技术需要“优化”数据集来监控模型优化的进展。

建模的另一个方面是变量产生/转换，如图 4 中步骤 **430** 所示。在这个处理过程中，对象是精确和领域知识的结合。原始数据值未必构成最好的模型变量，这是由于许多原因：例如，数据输入错误；非数字值；漏测值；和异常值。在运行建模逻辑之前，变量经常需

要再产生或转换以最好地使用收集的信息。为了避免开发数据、测试数据和确证数据之间的依赖性，所有的转换逻辑优选只来自于开发数据。

和转换如所期望的和/或如所需要的变量一起，建模过程包括变量选择步骤 440。其后，模型开发可包括培训模型 450 连同模型测试。随后可以是模型确证。

模型确证 460 的结果显示是否达到性能目标 470。如图 4 所示，如果性能目标已经达到，那么建模过程终止在步骤 480。假如性能目标没有达到，就需要模型的进一步开发。因此，图 4 的过程可以返回到步骤 430 以改变变量生成或转换，以便达到更好的性能。

### 实施例 3

#### 脉管斑块的分类

本实施例描述了本发明的使用统计分类装置来检测和分类脉管斑块的优选具体实施方式。首先，工作集中在利用成组模型建立系统，用于检测在活体外血管的 MR（磁共振）图像中动脉粥样硬化性斑的三种关键成分。该系统也检测动脉肌肉组织，当和脉管斑块和脂质检测系统结合时，该系统使得能够在图像中鉴定整个动脉和计算脉管斑块含量的估计值。这个系统是全自动的，并且在此实施例中，在检测和分析过程中，仅有的人类干预发生在从 MRI 设备收集原始磁共振数据期间。使用此系统，在脉管斑块成分分类中，成功率等于或好于人类放射学专家的表现。

在本实施例中，培训预测模型以鉴定三种组织类型：斑块、脂质、和肌肉。斑块检测器是使用确定硬脉管斑块的范例图像的标记来培训。脂质检测器是用脂质能被确定和标记的较小集的图像来培训。肌肉检测器用来从图像所示的血管的其它部分分离动脉壁组

织。可以开发另外的模型来检测钙化组织、血栓、和其他非病理性组织。伴随着动脉壁的正确鉴别，给定其他模型的输出值，就可以计算脉管中斑块含量估计值。

预测模型的目标是为图像的每个像素精确预测基本真值类别，其是基于在所述像素及其紧邻的周围环境处的 MRI 图像的特性。一旦完成图像处理步骤并且初始图像被转换成作为记录表示每个像素的柱状数据，则开始预测建模。每一个像素记录包含一个变量，该变量确定每一个像素的基本真值类别，并且超过 400 个另外的变量捕获衍生自图像处理步骤的特性。预测建模的任务是筛选这些成百的潜在变量和成千上万的变量排列，以提出最有预测性的组合。

在一优选的具体实施方式中，使用了人工神经网络 (ANN) 建模方法，其称作相关输入处理器网络 (RIPNet<sup>TM</sup>; Bates White, LLC, San Diego, CA; Perez-Amaral and White (2003), Oxford Bulletin of Economics and Statistics, vol. 65: 821- 838) ; 然而，也可以使用本领域技术人员所知的标准线性和非线性回归技术 (如线性和 logistic 回归，判定树，非参数回归 (如，使用神经网络或径向基函数)，Bayesian 网络建模，Fisher 判别分析，模糊逻辑系统等)。特别开发 RIPNet<sup>TM</sup> 来解决怎样鉴别有很多潜在变量的网络结构的问题，同时避免过度拟合。与神经网络估计有关的典型问题是在这些模型中体现的函数形式本质上是“太灵活”。标准的 ANN 方法规定了模型灵活度的水平，如果自动估计 (除非测试或交叉确证过程也被自动包括)，则其不但概括数据中的信号而且也概括噪音。这导致过度拟合，过度拟合是这样一种情境，其中模型并未很好地归纳出未包含在培训数据中的信息。过程如最佳停止规则已经广泛接受为一种方法，该方法用于在噪声拟合开始前停止网络培训过程 (尤其是最小二乘拟合算法)。这些程序处理故障现象，但不处理模型过度拟合的原因。在 ANN 中模型过度拟合基本是由模型规定中过度复杂性所引起，其完全类似于线性模型中可能遇到的过度拟合问题。当

在线性模型中遇到过度拟合时，一种解决办法是不去一般地修改最小二乘拟合程序（routine），而是通过去掉参数来简化模型的规定。另一种解决办法是使用交叉确证数据集来指示何时停止拟合。

RIPNet 体现了根本不同于神经网络估计的方法，其直接的目标是鉴别模型复杂性水平，这保证最好的在样本范围之外的预测性能而不特别修改拟合算法本身。在使用 RIPNet 方法产生模型的过程中有 5 个主要步骤，其在接着的结果部分更详细地加以讨论：（1）数据集生成、标记、和采样；（2）异常数据的检测；（3）变量预选和变换生成；（4）预测模型估计和变量选择；和（5）最终模型确证。

#### i. 数据集生成

从 MRI 图像的合并数据集开发模型以在逐个像素的基础上预测各个主要组织特性的存在。10 个标记图像产生了 112, 481 个可用的组织观测值（也就是像素）。结果说明特征检测的可扩展方法不依赖于血管几何形状的特性或图像分辨率，来获得临床相关的、可靠结果。

用于模型培训和估计的图像数据集包括 10 个活体外动脉切片，其代表了此课题可用的所有血管横断面图像。通过和组织结构直接对比来完成每一个动脉的斑块、肌肉、和脂质成分的标记。所有检查的图像包括为估计而被标记的硬脉管斑块的显著实例。在绝大多数图像中，肌肉被明显地鉴别和标记，但在一些图像中，例如，图 10 中的低质量图像，动脉之一提出了一个组织学的问题。仅三个图像包含了脂质的实例，而在图 5 中的图像具有最大的这样的实例。

从这些图像中提取的数据是像素形式的，这些像素被处理为分离的数据点。用于建模过程的目标变量是指示性变量，如果像素属

于目标类，则指示性变量为 1，否则为零。指示器是以图像的标记为基础的。在 MRI 图像中与每一个像素有关的是表明在 T1, T2, 和 PD 模态中强度的三个变量。这些强度的动态范围是从 0 到 255, 仅是整数值 (8 比特色深度)。大量地处理这些数据以生成大量附加变量, 其总结这样的情况如像素邻域中的平均强度, 以及其他更复杂的变换如局部结构度量。

因为样品安放在载物片上, 其不能生成有用的 MRI 信息, 所以大量像素因为未包含相关数据而被去除。此处, 这些像素被鉴别为那些 T1, T2, 和 PD 指示变量都同时为零的像素。这个程序是保守的, 并允许一些有随机噪声的像素进入数据集, 虽然这并不影响算法的实施。用这种方式每一个初始图像的 50% 以上被省略。

为了使结果最大化, RIPNet 程序优选在建模过程的多个阶段对数据进行拆分以便在整个实际实施中进行系统测试。由于这个原因, 一些图像被完全保留作为性能的测试。用于建模的数据集如下: 培训, 用来估计模型参数; 确证, 用于建模结果的交叉确证以基于样品外熵度量和伪-R-平方度量检验所选择变量的性能; 测试, 仅偶尔用于比较可选模型规范的相对性能, 这种数据集是通过 Data Miner's Reality Check™ 算法加以开发和应用 (White, H. (2000), *Econometrica*, vol. 68:1097-1126; 美国专利第 5,893,069 号和第 6,088,676 号), 这是因为确证数据集被大量地开发; 和支持, 其数据 (三种图像和超过 45,000 个观测结果) 被完全留在估计和确证过程之外以便向模型提供真实的实例。选择所有记录随机进入其相应的样品。

## ii. 异常数据的检测

开发了异常数据检测算法以确定数据中的异常值 (outliers)。这里, 异常检测器是聚类算法的形式, 其允许在数据中鉴别多变量



异常值。异常等同于离开（如通过 L1-标准测定的）其 k 最近邻域的记录。数据被分散在目标=0 样品和目标=1 样品之间以便可以相对于这些分离组对异常进行鉴定。在这种情况下，选择 k=10。这个过程通常是需要的，以鉴别那些对模型估计过程中可能有不平常影响的记录。然而，如果从成像设备获得的图像原始数据相对清楚，和大多数 MRI 数据的情况一样，则没有主要的异常值可以鉴别。

以下的表 1 列出一些异常变量的内容

表 1: 异常变量内容

斑块	脂质	肌肉
PD 平均数 3 像素半径	T1 最小值 3 像素半径	T2 最小值 3 像素半径
T2 平均数 3 像素半径	T2 最小值 3 像素半径	T1 最小值 5 像素半径
T2 中位数 3 像素半径	T1 最小值 4 像素半径	T2 平均数 5 像素半径
T1 平均数 5 像素半径	T1 最小值 5 像素半径	T2 最小值 5 像素半径
K 均值变量	T2 最小值 5 像素半径	K 均值变量

异常数据检测引擎也生成异常变量。该异常变量将每一个观测值（相对于目标=0 的样品和相对于目标=1 的样品）的两个距离度量转换成似然比统计量。在输入变量的变换中这个统计量把相对距离嵌入目标样品，在某些情况下，其是有效的预测值。每个模型的异常变量由多达 5 个连续的输入变量组成，如以上表 1 所示。

### iii. 变量预选和变换生成

接下来，需进行附加的阶段，即变量变换生成和初步消除非预测变量，以缩减数据集的大小。此阶段生成的变换包括以下变换（对于输入数据集的所有变量）：群变换，其是被分为十分位数收集箱（bins）的单变量连续变量，然后通过聚类算法来组合十分位数收集箱以获得最少数量的收集箱而没有明显降低预测性能；叉积，其是单变量连续和离散变量，这些变量相互影响并使用上述的聚类算法被分为基于收集箱（binned）的范畴变量；和 $\beta$ 变换，这些 $\beta$ 变换

是灵活的函数形式，其基于对数据进行 $\beta$ 分布函数拟合和计算似然比。

对到现在为止生成的所有变量做性能测试，其是基于目标变量并利用样品之外的伪-R-平方统计量。直接的熵计算对比了在给定因变量状态下独立变量的分布，其然后被总结在伪-R-平方统计量中用于确证样本。在小样本中，伪-R-平方统计量不是限制在0和1之间（因为范围和估计样本的范围不正好相同）。十个最大的预测群变换变量和十个最大的预测叉积变量保留在数据集中并传递到模型评估程序。仅前五位的预测 $\beta$ 变换变量被保留下来。有低或负的单变量或双变量伪-R-平方统计量的变量也从潜在的候选变量库中被永久地去除。这些变量变换列在以下的表2中。

表2：变量变换

群变换	叉积	$\beta$ 变换
异常	K 均值 × 异常	异常
K 均值	T2 平均数 3 像素半径 × 异常	K 均值
T2 平均数 3 像素半径	PD 3 像素预处理 × 异常	T2 平均数 3 像素半径
T2 中位数 3 像素半径	T2 中位数 3 像素半径 × 异常	T2 中位数 3 像素半径
PD 3 像素预处理	T2 平均数 3 像素半径 × K 均值	pd 3 像素预处理
T1 平均数 5 像素半径	T2 中位数 3 像素半径 × K 均值	
T1 中位数 5 像素半径	PD 3 像素预处理 × K 均值	
T2 中位数 5 像素半径	PD 3 像素预处理 × T2 平均数 3 像素半径	
黄色空间	T2 中位数 3 像素半径 × T2 平均数 3 像素半径	
PD 5 像素预处理	PD 3 像素预处理 × T2 中位数 3 像素半径	

#### iv. 模型估计和变量选择

正如大多数模式识别实施例的情况一样，存在远多于在预测模型中实际上可能容纳的潜在候选变量要包括，这引起了一些重大的危险。一是潜在有用的候选变量被忽略，仅仅因为有太多的变量要被评价。二是如果使用用于评价和包含变量的系统程序，它能引起过度拟合。最后，很多候选变量可能是冗余的，这能引起估计程序的问题。例如，T1 模态的平均值是从 3-像素半径到 9-像素半径的邻域范围内获得，并且都包括作为候选变量。

在此实施例中使用的 RIPNet<sup>TM</sup> 程序处理这些危险，其是通过结合典型 ANN 的理论非线性曲线拟合能力和分级技术的稳定性。这种对输入变量的非线性组合和变换的搜索可以随后用在标准最大似然分对数 (logit) 模型。RIPNet<sup>TM</sup> 包含用于变量生成 (网络节点)、变量测试、和模型估计的算法。

有两个输入和一个输出的典型的单隐藏层前馈网络可以有以下结构：

其中  $f(\cdot)$  和  $g(\cdot)$  是所谓的压碎 (s-型) 函数。这些压碎函数传送 ANN 的能力 (power)，因为它们表现出依赖于参数  $\beta$  和  $\gamma$  设定的一些不同的行为。这类设定的实例包括：反函数；对数函数；指数函数；和阈值函数。

RIPNet 建模策略以如下面的函数形式来开始：

$$y = g(y_1 f(x_1 \beta_{11} + x_2 \beta_{12}) + y_2 f(x_1 \beta_{21} + x_2 \beta_{22}))$$

其丰度和非线性源自于函数  $f_i(x_1, x_2)$ 。RIPNet 算法的主要贡献之一是用于生成模拟的网络节点  $f_i(x_1, x_2)$  的高产率方法。尽管它外表简单，这种形式的模型可以用来紧密地接近传统 ANN 的性能。

在相对易处理模型的类型中的节点选择是处理过程中的下一步骤。正如变量预筛选步骤一样，节点选择是基于使用确证样本来检查样本之外的性能。按照候选节点在预测模型中的经确证的预测性能的次序，把候选节点引入模型中。在选择程序中以两种方式处理冗余。首先，当附加的节点引入模型中时，它们被正交化以便去除冗余成分。测试多重阈值以便能够确定节点正交性的最优水平。第二，选择冗余的阈值以致仅小于它们的方差 5% 的节点 (由模型中的其它节点解释) 可以引入。

## v. 模型确证

最终模型确证步骤用来确保无论哪个模型被选为最终模型，它都好于较简单的基准或其他候选模型。Data Miner's Reality Check™ (DMRC) 再次用来以此方式测试模型。这项技术利用样品外的预测和自举分配，来为测试模型具有和基准模型相同性能的假说生成有效的 p-值。低 p 值表明测试模型显示出明显更好的性能。

## v. 结果

相对于目前应用的主导技术，本文描述的预测模型具有明显改进的预测性能。图 9 总结了以上描述的模型的性能。在图中，表 1 总结了 RIPNet 模型的性能，其是基于两种不同的统计方法：最大 Kolmogorov-Smirnov 统计量；和 Gini 系数，其每一个测定 ROC (局部运算 (操作) 特性) 曲线的情况。用基于 K-均值法的模型作为对比的基础。这些结果说明 RIPNet 模型在绝对意义上普遍地优于 K-均值法 25%到 30%。给定假阳性的水平，这转化到 50%更高的真阳性率。ROC 曲线 (从其获得这些结果) 示于图 9 中。

除统计性能度量外，以图像形式 (图 8) 显示的组合模型结果也支持这一结论。在图 8 中，对于每一个图像，标记基本真值表示在左图片上，而对于相同图像的模型预测则表示在右图片上。在图中，在标记图像中肌肉显示为粉红色，而在模型结果中是红色。在标记图像中脂质显示为白色，在模型结果中是蓝色，而脉管斑块在两个图像中都表示为黄色。每个像素被指定类别，其依赖于哪个模型为所述像素生成最高的概率。对于所有的模型预测，低于 30% 概率的像素被编码为空白。

从预测模型生成的并示于图 8A 中的图像提供了本发明的预测模型能力的清楚的实例。特别是，在初始的、基本真值的、标记图

像中像素标记斑块和那些通过计算模型标记的斑块之间有高度的一致性。同样，也很好鉴别了肌肉。在图 8B 中的图像说明本发明的模型不仅对硬脉管斑块，而且对有脂质成分的脉管斑块也有检测能力。肌肉区未被均匀地加以鉴别，但这精确地发生在这样的区域：初始图像被伪像干扰的区域，以及肌肉壁较薄的区域。有趣地，沿着包围脂质核心（示于此图中）的纤维盖的区域，有一些假阳性来自肌肉模型。

在此项分析中包括 400 个以上的变量，并且从这些变量产生了数千的网络节点。在所有这些节点中，为每一个模型选择了 140 个节点。这通常被认为是很大的数量，但对于数据集中的大量观测值而言则不是。以文的表 3 说明了为每个模型选择的几个顶级节点。这些节点的检查说明了相对于其他方法而言使用自动化技术的益处。例如，脉管斑块模型主要包含 T1 和 PD 变量，而对于绝大多数包括的变量，对所有其它变量选择了 7-像素邻域度量。这些类型的选择几乎不可能在没有巨大努力的情况下人工再生。同样，对于人眼来说，变量的许多组合是不明显的。例如，脂质模型的顶级节点是最大 4 像素邻域（对于 T2）、PD 的离散余弦变换、和异常变量的线性组合，对其现今没有清楚的解释。当然没有一个常规的启发式方法会发现这些。

表 3: 顶级 (top) 有关输入

脉管斑块模型		
分类	类型	说明
1	节点	K 均值
2	节点	7 像素半径的 T1 平均值
3	节点	7 像素半径的 PD 中位数
4	节点	3 像素半径的 T1 最小值
5	节点	7 像素半径的 T1 最小值
6	节点	7 像素半径的 PD DCT

肌肉模型		
分类	类型	说明
1	节点	异常
2	节点	异常 T2 变量 4 像素半径
3	节点	异常 PD 偏移 (Skew) 3 像素半径
4	隐藏单位 (unit)	T1 预处理 3 像素半径 T2 预处理 5 像素半径 CR 7 像素半径的 PD DCT 异常变量

脂质模型		
分类	类型	说明
1	隐藏单位	4 像素半径的 T2 最大值 PD DCT 5 像素半径 异常
2	隐藏单位	3 像素半径的 T1 平均值 3 像素半径的 PD 最大值 黄色-CMYK 变换 5 像素半径的 PD DCT 异常变量
3	隐藏单位	3 像素半径的 T2 范围 4 像素半径的 PD 方差 5 像素半径的 PD IQR 6 像素半径的 PD DCT 异常

重要的是还要注意这些模型的高度非线性特性。异常变量表现为最预测的变量之一的事实用于着重说明这一事实。隐藏单位、群变换、 $\beta$ 变换等都是输入数据的非线性变换，这些输入数据在这些模型中表现为顶级变量。

为了避免过度拟合，采取了预防措施。对从建模过程中完全被保留的一些图像进行测试。如图 10 所示，模型在分离肌肉和脉管斑块中相当好地完成了任务。肌肉模型跟踪动脉壁的基本轮廓，并且该模型很好地识别标记的脉管斑块区域。当这种图像要求专业放射学家首先标记时，它促进了挑战的和最终成功的测试。

## 实施例 4

### 3-D 血管模型

本实施例描述了一种生成血管三维模型的优选方法，该血管已经使用医学成像仪成像。具体地说，图 13 显示了在颈动脉杈区域的颈动脉的 3-D 透视图。在此图中，动脉内壁（1920）代表腔的边界。斑块（1910）位于动脉壁内表面和动脉外表面（未示出）之间。这个模型来源于十一个体内 MRI 组织切片并仅利用 T1 模式。为了生成该模型，使用以下步骤来转换组织切片图像。首先，每一个切片的数据通过低通滤波器（例如，自适应 Wiener 滤波器）。每个切片的腔中心位置是基于来自前一个切片的腔形心的位置来评估。然后在集中于估计的腔位置后，裁剪每个图像。接着，进行阈值强度的线性搜索以显示接近估计形心的腔区域。在确证该腔有包括区域和离心率在内的必要的形态特征后，重新评估腔形心的位置。接着，进行组织分割以确定靠近腔形心的脂质特征。

如将明了的，依据切片是高于还是低于颈动脉杈，稍微修改前述的过程。当算法追踪两个腔（即，在颈动脉杈上方的切片中）时，就使用两个腔形心的几何平均数。然后为了补偿轴向偏移，得到的切片被重新定中心。也可以包括附加的其它组织信息，其包括关于下述的信息：肌肉，外膜，和脉管斑块成分如脂质，出血，纤维斑块、和钙。得到的模型将便于脉管斑块大小、体积、和成分的可视化和自动量化。

\* \* \*

本文描述和要求的所有方法，系统和制造物品可以根据本说明书并且不需要过多实验的情况下就可以制造和执行。虽然本发明的方法，系统、和计算机程序产品已通过优选具体实施方式和可选的特性加以描述，但是，对于本领域技术人员来说，显而易见的是，

可以对本文描述的方法以及所述方法的步骤或步骤顺序进行改进和变化，而不偏离本发明的精神和范围。更具体地说，显而易见的是，不同的算法，软件、和数据可以适用于脉管斑块的自动探测和分析。所有这样等效的或相似的适应，修饰，修改，和对于本领域技术人员来说显而易见的替代物被认为是在如所附权利要求所定义的本发明的精神和范围内。

本文已对本发明进行了广泛且一般性的描述。每一个属于本一般披露内容的较小的种类和亚属分类也形成本发明的部分。这包括本发明的一般描述，附带条件或负限制为从种类中去除任何主题，不管已删除的材料是否在本文中被明确地描述。

本文说明性描述的发明可以适当地在不存在任何未在本文中特别披露为基本的要素的情况下进行实施。已经使用的术语和措辞是以描述性而不是限制性的术语来使用的，因此在使用这样的术语和措辞时并不排除示出和描述的任何现已存在或以后开发的等价特性或其部分，而是要认识到，在本发明权利要求范围内各种改进是可能的。同样，术语“包含”，“包括”，“含有”等等是被广泛地理解和没有限制的。必须注意的是，如在本文中和在所附权利要求中所使用的，单一形式“一”，“一种”和“该”包含复数含义，除非上下文另有明确说明。

所有在本说明书中提及的专利，专利申请，和出版物显示与本发明有关的技术领域中的普通技术人员的水平。所有的专利、专利申请、以及出版物以引用方式结合于本文中，其是为了所有目的和达到相同程度，好像每一个单独的专利、专利申请、或出版物都明确和单独显示成通过引用被整合。



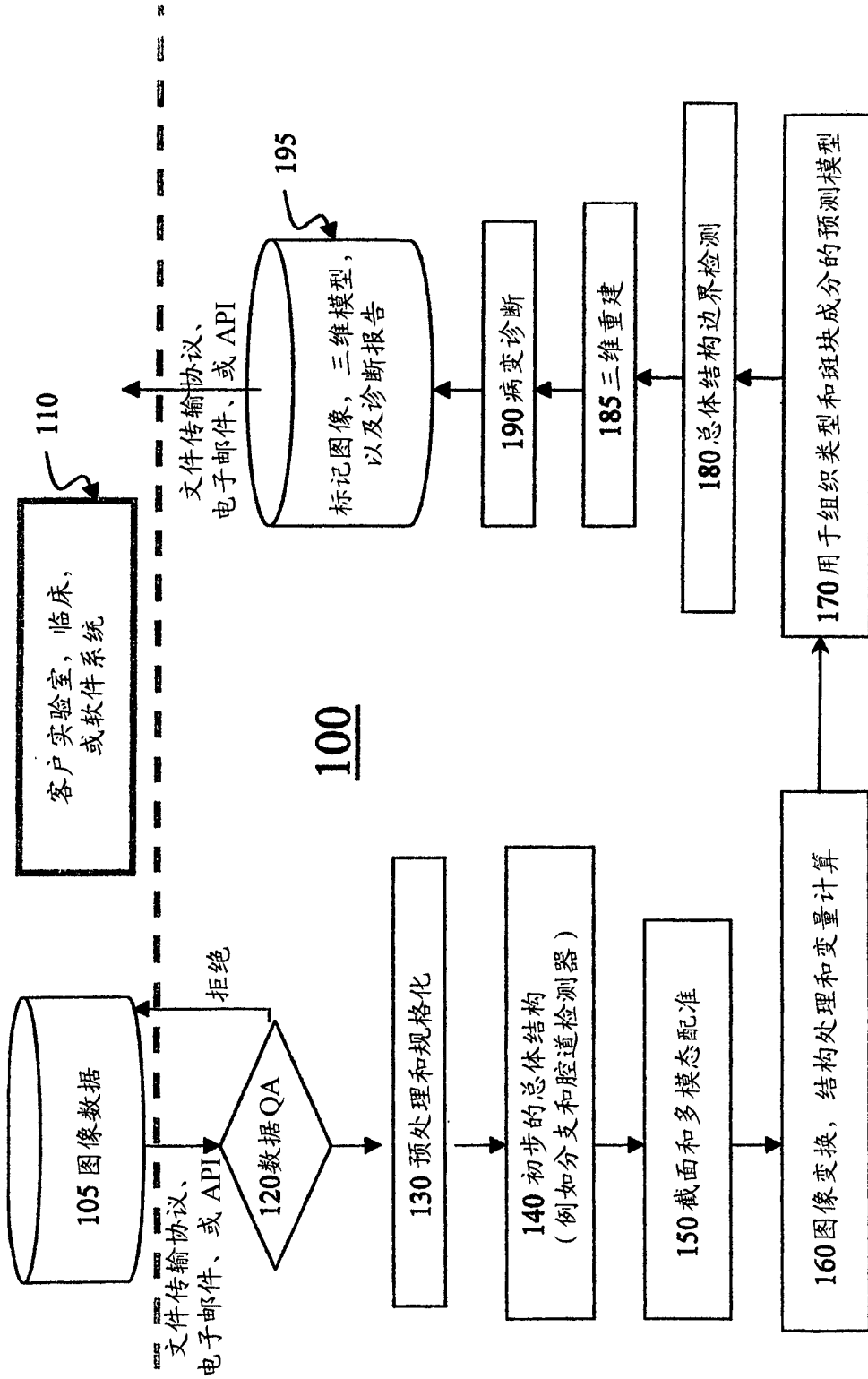


图 1

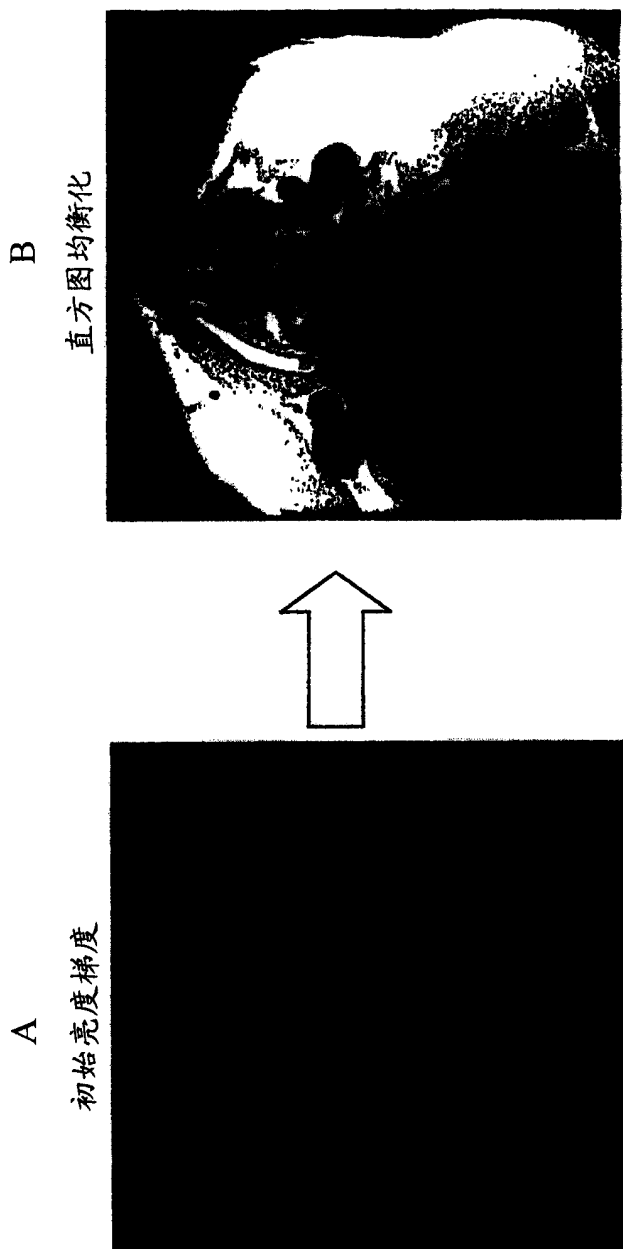


图 2

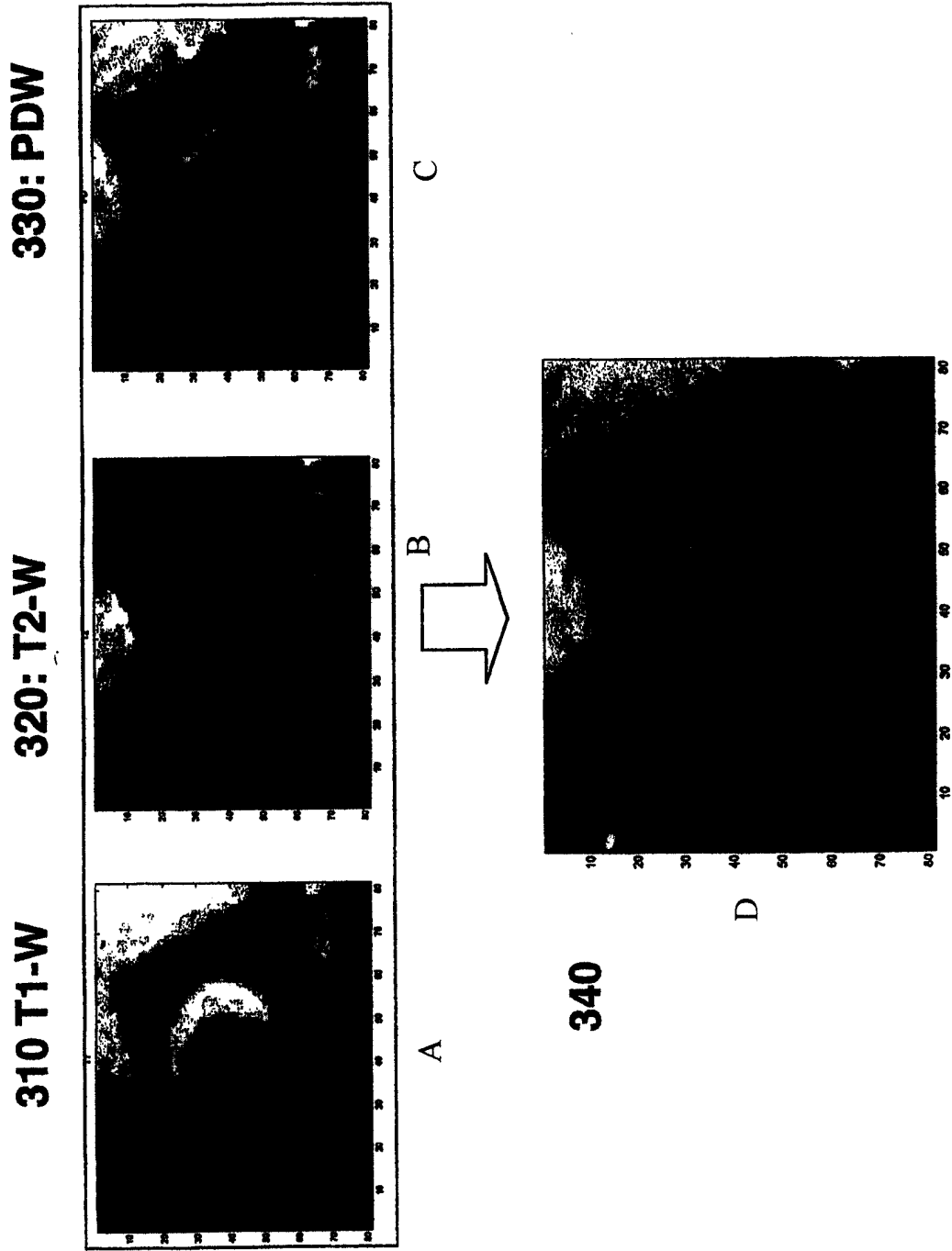


图 3

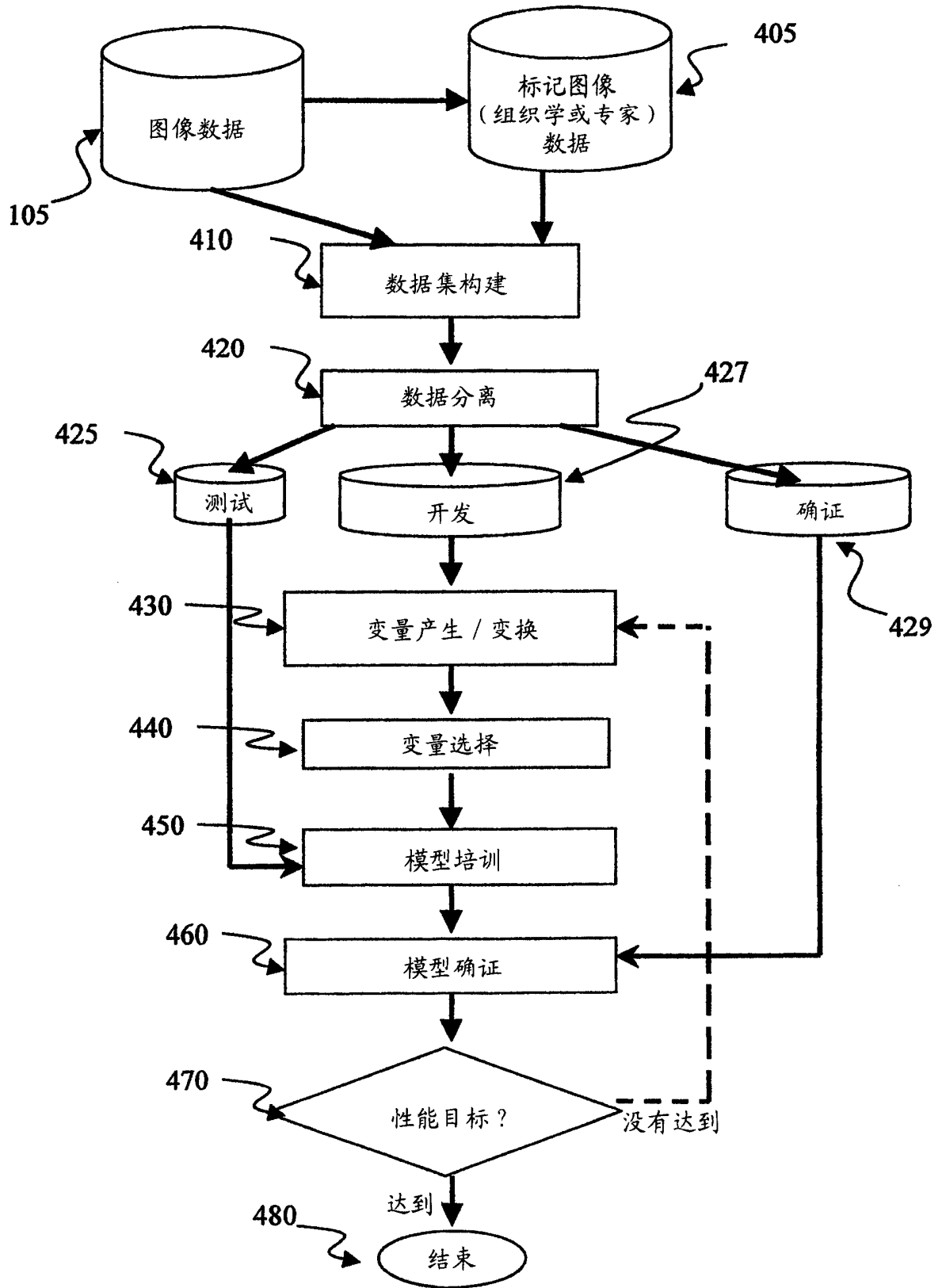


图 4

A. MRI 图像

510 PDW  
(绿色)

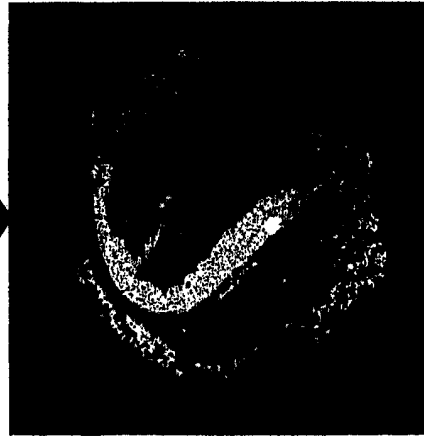
520 T1 W  
(红色)

530 T2 W  
(蓝色)



B. 三色图像

540 伪彩色图像



550 C. 组织病理学 (确证)



图 5

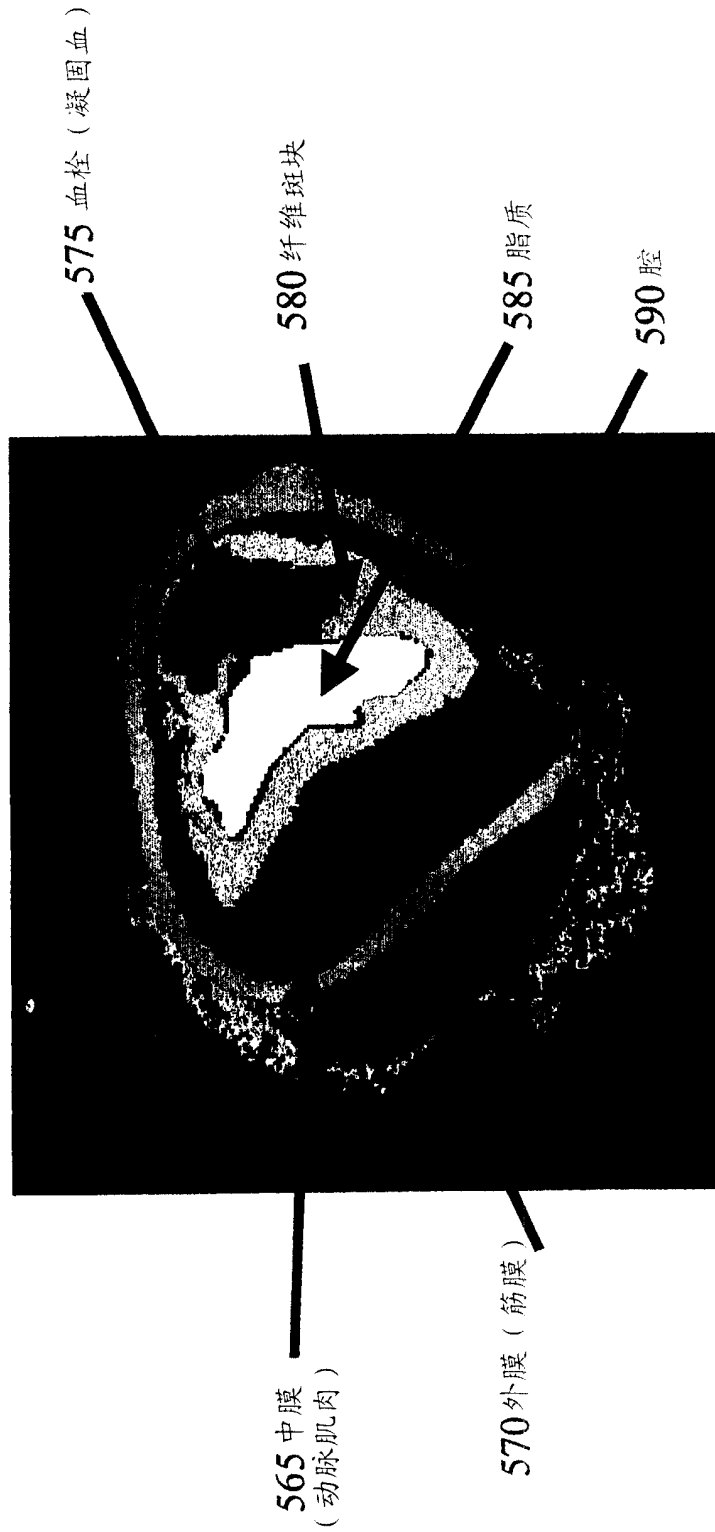


图 5 (Cont.)

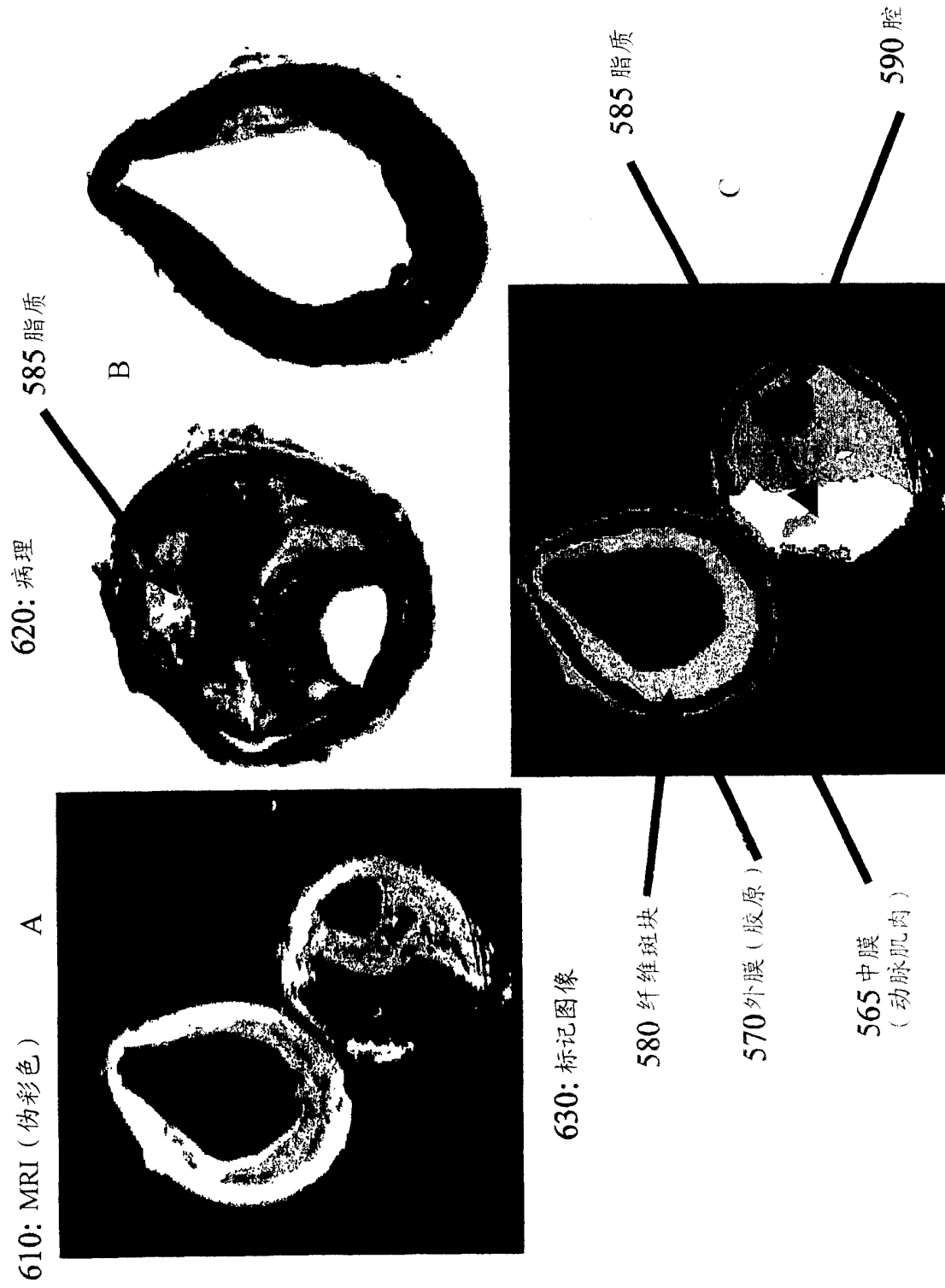


图6

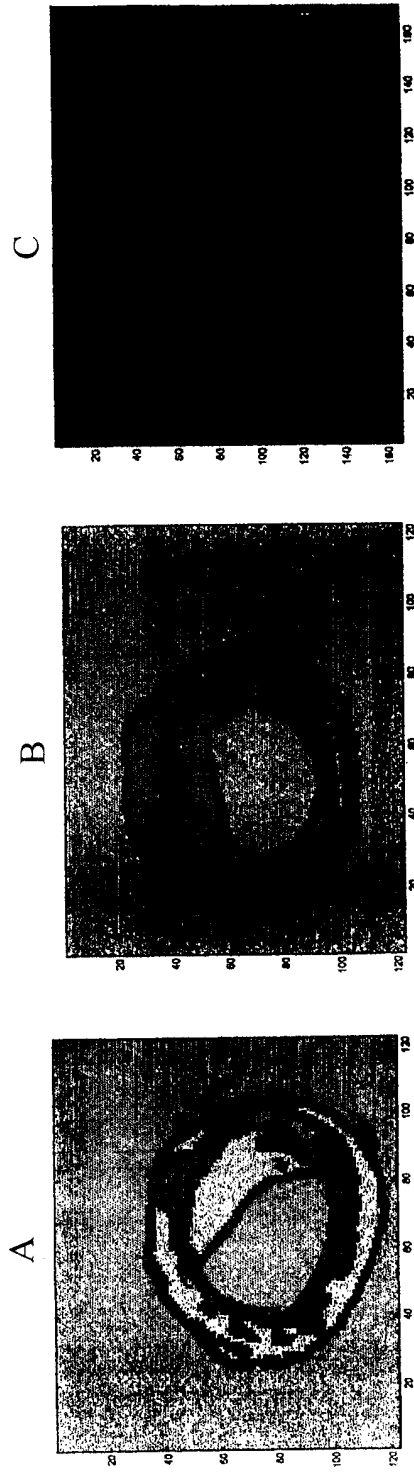
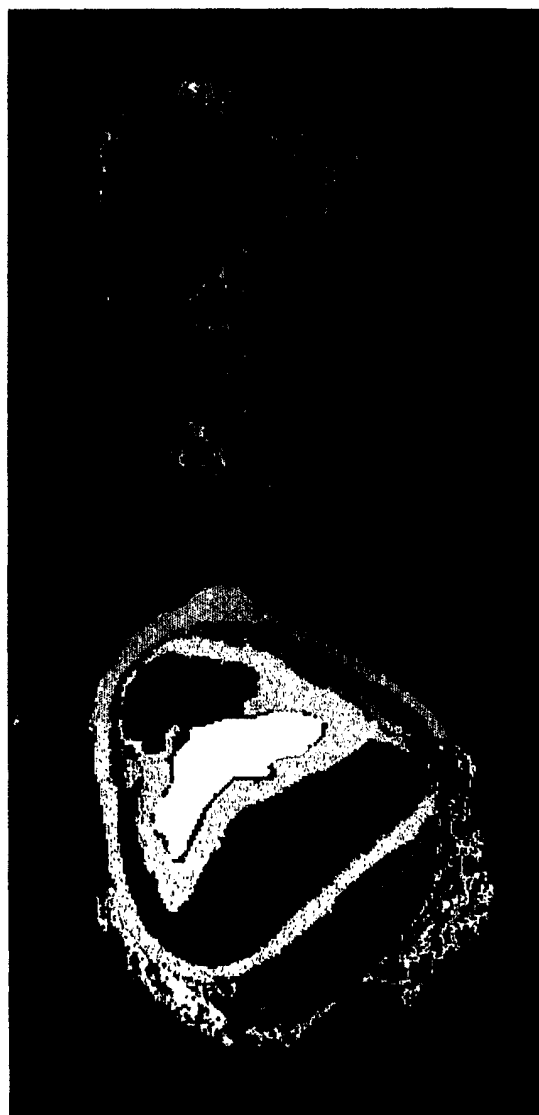


图7





A



B

图 8

表 A: 模型性能

模型	最大 KS		Gini	
	Ripnet	K-均值	Ripnet	K-均值
斑块	76.1%	57.4%	83.8%	67.1%
肌肉	82.5%	57.3%	89.7%	56.6%
脂质	91.4%	58.0%	94.5%	55.3%
平均	<b>83.3%</b>	<b>57.6%</b>	<b>89.3%</b>	<b>59.7%</b>

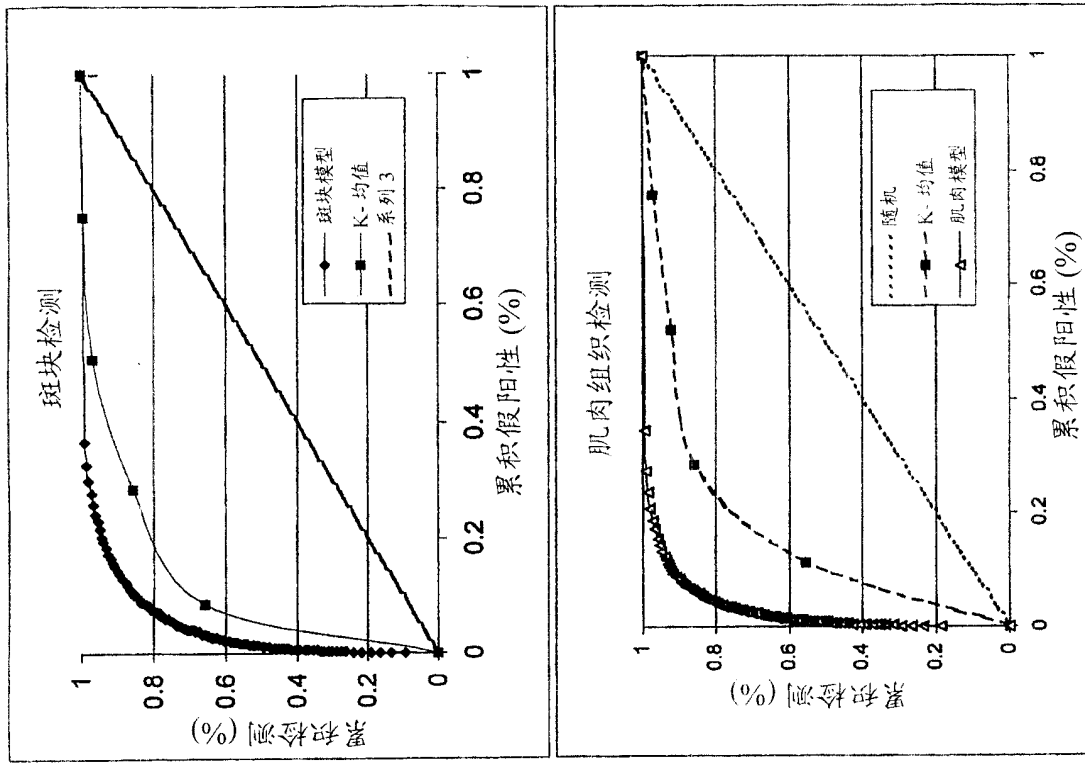


图 9

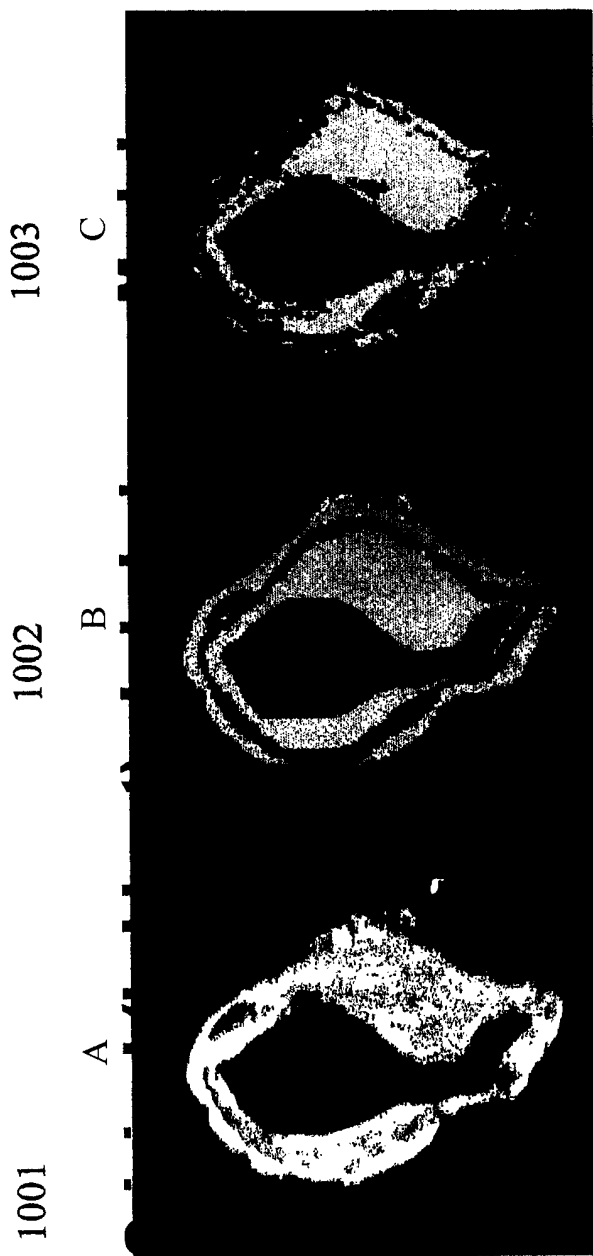


图 10

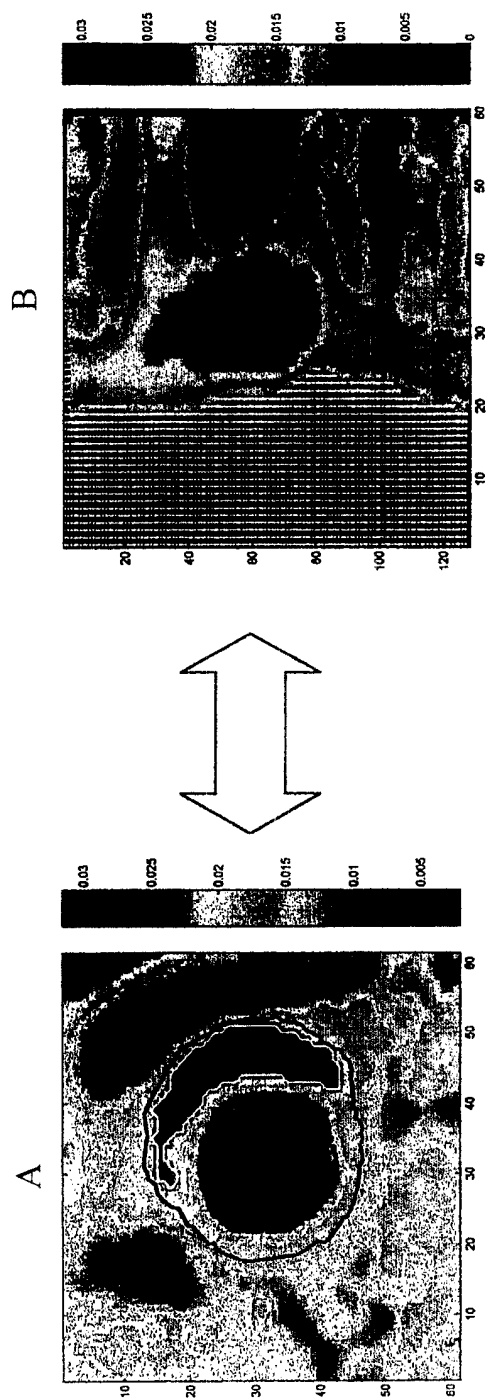


图 11

使用梯度向量流的肌肉壁示踪

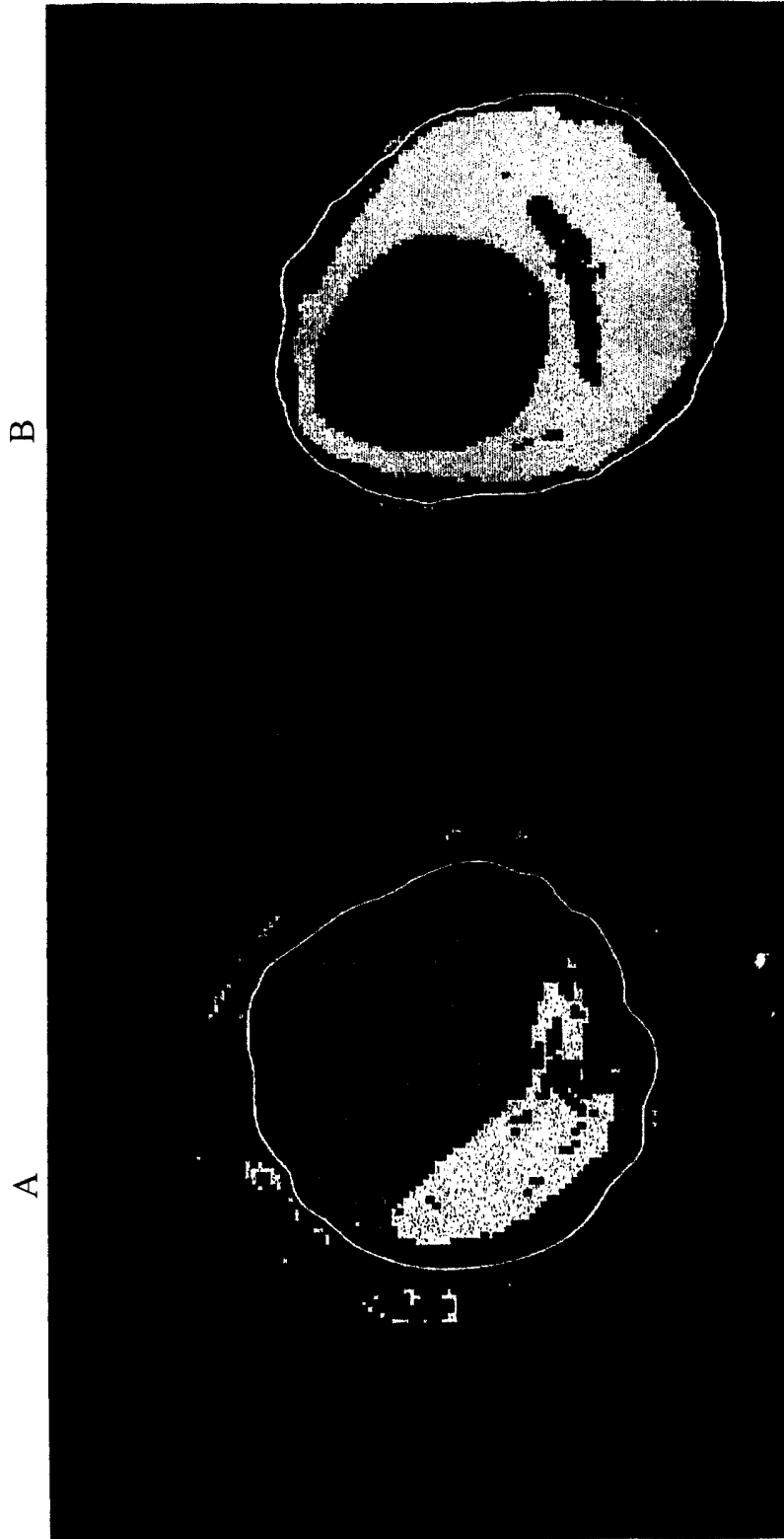


图 12

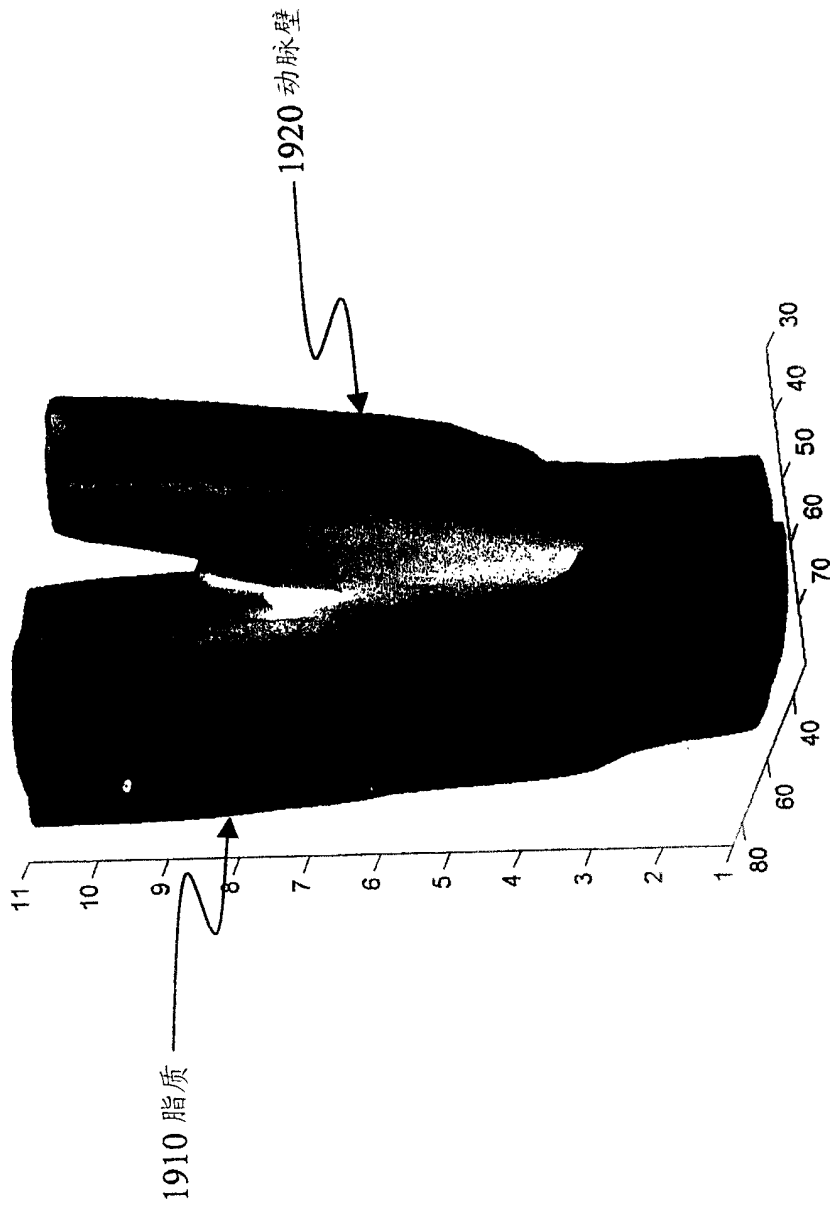


图 13