



(12) 发明专利申请

(10) 申请公布号 CN 118541706 A

(43) 申请公布日 2024. 08. 23

(21) 申请号 202280086521.6

(22) 申请日 2022.10.19

(30) 优先权数据

21205556.0 2021.10.29 EP

(85) PCT国际申请进入国家阶段日

2024.06.27

(86) PCT国际申请的申请数据

PCT/GB2022/052663 2022.10.19

(87) PCT国际申请的公布数据

W02023/073346 EN 2023.05.04

(71) 申请人 帕纳基亚科技有限公司

地址 英国

(72) 发明人 M·P·拉哈尔加 李修思

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

专利代理师 王伟楠 何方

(51) Int.Cl.

G06N 20/20 (2019.01)

G06F 18/25 (2023.01)

G06N 3/09 (2023.01)

G06N 3/0464 (2023.01)

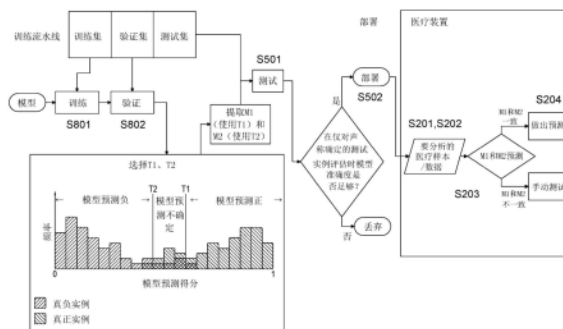
权利要求书2页 说明书28页 附图24页

(54) 发明名称

分类方法

(57) 摘要

一种计算机实现的方法,包括:获得包括与第一类对应的第一输入数据的第一数据集;获得包括与第一类对应的第二输入数据的第二数据集;使用第一数据集训练至少一个分类器;将来自第二数据集的第二输入数据输入至至少一个分类器,并且提供包括第一分类和第二分类的分类模型,其中,第一分类与第二分类相比预测与第一类对应的第二输入数据处于第一类中的比例更大。



1. 一种计算机实现的方法,包括:

获得包括与第一类对应的第一输入数据的第一数据集;

获得包括与所述第一类对应的第二输入数据的第二数据集;

使用所述第一数据集训练至少一个分类器;

将来自所述第二数据集的所述第二输入数据输入至所述至少一个分类器,并且提供包括第一分类和第二分类的分类模型,其中,所述第一分类与所述第二分类相比预测与所述第一类对应的所述第二输入数据处于所述第一类中的比例更大。

2. 根据权利要求1所述的方法,其中,所述第一分类包括应用第一分类器和使用第一预测阈值,并且所述第二分类包括应用所述第一分类器和使用第二预测阈值,其中,确定所述第一预测阈值包括确定所述第一分类器针对来自所述第二数据集的与所述第一类对应的输入的最高输出值,并且将所述第一预测阈值设置为小于或等于所述最高输出值,并且其中,确定所述第二预测阈值包括确定针对来自所述第二数据集的与第二类对应的输入的最低输出值,并且将所述第二预测阈值设置为大于或等于所述最低输出值。

3. 根据权利要求1所述的方法,其中,所述至少一个分类器包括至少两个分类器,并且其中,提供所述分类模型包括:确定所述至少两个分类器中的第一分类器,其中,在所述至少两个分类器中所述第一分类器预测与所述第一类对应的所述第二输入数据处于所述第一类中的比例最大;以及确定所述至少两个分类器中的第二分类器,其中,在所述至少两个分类器中所述第二分类器预测与第二类对应的所述第二输入数据处于所述第二类中的比例最大。

4. 根据权利要求2或3所述的方法,其中,提供所述分类模型包括确定第一预测阈值和第二预测阈值,其中,确定所述第一预测阈值和所述第二预测阈值包括:获得所述第一预测阈值的值和所述第二预测阈值的值的多个可能组合;确定使用所述可能组合中的每一个组合的分类模型的准确度;以及选择具有高于预定准确度值的准确度的组合。

5. 一种计算机实现的方法,包括:

获得包括与第一类对应的第一输入数据的第一数据集;

获得包括与所述第一类对应的第二输入数据的第二数据集;

使用所述第一数据集训练分类器;

将来自所述第二数据集的所述第二输入数据输入至至少一个分类器,并且提供包括第一分类的分类模型,其中,提供包括所述第一分类的所述分类模型包括确定与所述第一类对应的预测阈值。

6. 根据权利要求1至5中任一项所述的方法,其中,所述第一数据集包括与第一群体对应的第一输入数据,并且其中,所述第二数据集包括与第二群体对应的第二输入数据。

7. 一种计算机实现的方法,包括:

获得输入数据;

将所述输入数据输入至分类模型,所述分类模型应用第一分类和第二分类,其中,所述第一分类被配置成与所述第二分类相比正确地预测与第一类对应的输入数据的比例更大;基于所述第一分类和所述第二分类的输出确定所述输入数据的分类预测。

8. 根据权利要求1或7所述的方法,其中,所述第一分类包括应用第一分类器和使用第一预测阈值,并且所述第二分类包括应用所述第一分类器和使用第二预测阈值,其中,所述

第一预测阈值高于所述第二预测阈值。

9. 根据权利要求7所述的方法, 其中, 所述第一分类包括应用第一分类器, 并且所述第二分类包括应用第二分类器。

10. 根据权利要求9所述的方法, 其中, 所述第一分类使用第一预测阈值, 并且所述第二分类使用第二预测阈值, 其中, 所述第一预测阈值高于所述第二预测阈值。

11. 根据权利要求7至9中任一项所述的方法, 其中, 所述输入数据包括组织图像, 并且其中, 确定所述分类包括确定与医疗诊断有关的信息。

12. 一种计算机实现的分类方法, 包括:

获得输入数据;

将所述输入数据输入至分类模型, 所述分类模型应用第一分类, 其中, 所述第一分类包括应用第一分类器和使用与第一类对应的预测阈值;

基于所述第一分类的输出来确定所述第一类的分类预测。

13. 一种分类系统, 包括:

一个或多个处理器, 所述一个或多个处理器被配置成执行根据权利要求7至12中任一项所述的方法。

14. 一种分类系统, 包括根据权利要求1至7中任一项所述的方法训练的分类模型。

15. 一种载体介质, 包括计算机可读代码, 所述计算机可读代码被配置成使计算机执行根据权利要求1至12中任一项所述的方法。

分类方法

技术领域

[0001] 本公开内容涉及分类方法、分类系统和训练方法。

背景技术

[0002] 分类器模型在广泛的领域中都具有应用。例如,分类器模型可以用于各种医疗应用中。分类器模型可以用于通过将例如生理测量值作为输入来提供医疗诊断。在另一示例中,分类器模型可以用于分析医学图像数据,例如用于预测输入组织图像是否包含癌细胞,或者用于预测特定生物标志物是否与组织图像相关联。

[0003] 这样的分类器模型可以使用训练数据集进行训练。例如,训练数据集可以包括大量示例输入以及标记的分类预测。然而,使用针对训练数据集已经建模的分类器模型的分方法在针对新输入数据部署时将具有一定水平的不准确度。分类方法的准确度可以用正确预测的数目占通过该方法进行的全部预测的数目的比例来测量。例如,在针对数据集中的每个输入已知正确的预测的情况下,这可以针对测试数据集进行评估。

[0004] 在一些使用实例例如医疗使用实例中,需要高的准确度。可以通过提高分类器模型性能来提高分类方法的准确度。这样的提高可以通过针对例如新的训练数据集进一步训练分类器模型来实现。然而,在一些情况下,例如在不能容易地获得另外的训练数据的情况下,不可能进一步提高分类器模型性能。

附图说明

[0005] 现在将参照附图描述根据非限制性实施方式的系统和方法,在附图中:

[0006] 图1示出了根据实施方式的分类方法的示意图;

[0007] 图2(a)是针对数据集的分类器模型输出的示意图;

[0008] 图2(b)示出了根据第一示例的方法的示意图;

[0009] 图3示出了根据第一示例的方法的示意图;

[0010] 图4示出了卷积神经网络的示例;

[0011] 图5是示出了根据第一示例的训练流水线和经训练的分类系统的部署的示意图;

[0012] 图6是示出第一预测阈值的示意图;

[0013] 图7是示出第二预测阈值的示意图;

[0014] 图8(a)是根据第一示例的训练分类系统的方法的示意图;

[0015] 图8(b)是根据第一示例的训练分类系统的方法的示意图;

[0016] 图9示出了从完全灵敏度到完全特异度的过渡的示意图;

[0017] 图10(a)是根据第一示例的修改的训练分类系统的方法的示意图;

[0018] 图10(b)示出了优化处理的示例;

[0019] 图10(c)示出了根据阈值组合计算的准确度值;

[0020] 图11(a)示出了根据第二示例的分类方法;

[0021] 图11(b)示出了包括输入数据的方法的示意图;

- [0022] 图12(a)是根据第二示例的训练分类系统的方法的示意图；
- [0023] 图12(b)是示出根据第二示例的训练流水线和经训练的分类系统的部署的示意图；
- [0024] 图12(c)是根据第二示例的修改的训练分类器的方法的示意图；
- [0025] 图13是根据第三示例的分类方法的示意图；
- [0026] 图14是示例分类方法的示意图；
- [0027] 图15示出了示例分类系统的示意图；
- [0028] 图16示出了根据第四示例的分类方法。

具体实施方式

- [0029] 根据第一方面,提供了一种计算机实现的方法,包括:
- [0030] 获得包括与第一类对应的第一输入数据的第一数据集；
- [0031] 获得包括与第一类对应的第二输入数据的第二数据集；
- [0032] 使用第一数据集训练至少一个分类器；
- [0033] 将来自第二数据集的第二输入数据输入至至少一个分类器,并且提供包括第一分类和第二分类的分类模型,其中,第一分类与第二分类相比预测与第一类对应的第二输入数据处于第一类中的比例更大。
- [0034] 在示例中,第二分类与第一分类相比预测与第二类对应的第二输入数据处于第二类中的比例更大。
- [0035] 在示例中,第一类是负类,并且第二类是正类。
- [0036] 在示例中,第一分类包括应用第一分类器和使用第一预测阈值,并且第二分类包括应用第一分类器和使用第二预测阈值。第一预测阈值高于第二预测阈值。
- [0037] 在示例中,第一分类器是二元分类器。确定第一预测阈值可以包括确定二元分类器针对来自第二数据集的与第一类对应的输入的最高输出值,并且将第一预测阈值设置为小于或等于最高输出值,并且其中,确定第二预测阈值包括确定针对来自第二数据集的与第二类对应的输入的最低输出值,并且将第二预测阈值设置为大于或等于最低输出值。
- [0038] 在示例中,至少一个分类器包括至少两个分类器,并且提供分类模型包括确定至少两个分类器中的第一分类器,其中,在至少两个分类器中第一分类器预测与第一类对应的第二输入数据处于第一类中的比例最大,并且确定至少两个分类器中的第二分类器,其中,在至少两个分类器中第二分类器预测与第二类对应的第二输入数据处于第二类中的比例最大。
- [0039] 在示例中,提供分类模型包括:确定第一预测阈值和第二预测阈值,其中,确定第一预测阈值和第二预测阈值包括:获得第一预测阈值的值和第二预测阈值的值的多个可能组合;确定使用可能组合中的每一个组合的分类模型的准确度;以及选择具有高于预定准确度值的准确度的组合。
- [0040] 在示例中,第一数据集包括用于第一输入数据的分类标签,并且第二数据集包括用于第二输入数据的分类标签。
- [0041] 在示例中,第一分类器和第二分类器对应于不同的模型架构以及/或者对应于基于不同的输入数据和/或使用不同的超参数更新的参数。

[0042] 在示例中,第一数据集包括与第一群体对应的第一输入数据,并且第二数据集包括与第二群体对应的第二输入数据。

[0043] 根据另一方面,提供了一种计算机实现的方法,包括:

[0044] 获得包括与第一类对应的第一输入数据的第一数据集;

[0045] 获得包括与第一类对应的第二输入数据的第二数据集;

[0046] 使用第一数据集训练分类器;

[0047] 将来自第二数据集的第二输入数据输入至分类器,并且提供包括第一分类的分类模型,其中,提供包括第一分类的分类模型包括确定与第一类对应的预测阈值。

[0048] 在示例中,确定预测阈值包括确定分类器针对来自第二数据集的与第一类对应的输入的最高输出值,并且将预测阈值设置为小于或等于最高输出值。

[0049] 在示例中,确定预测阈值包括确定针对来自第二数据集的与第一类对应的输入的最低输出值,并且将预测阈值设置为大于或等于最低输出值。

[0050] 在示例中,确定预测阈值包括:获得预测阈值的多个可能值;使用可能值中的每一个值确定分类模型的得分;以及选择具有高于预定得分的得分的值。得分可以是与第一类不对应的数据被分类为处于第一类中的比例的度量。多个可能值可以包括在分类器针对来自第二数据集的与第一类对应的输入的最高输出值与分类器针对来自第二数据集的与第一类对应的输入的最低输出值之间的值集。

[0051] 根据另一方面,提供了一种计算机实现的方法,包括:

[0052] 获得输入数据;

[0053] 将输入数据输入至分类模型,分类模型应用第一分类和第二分类,其中,第一分类被配置成与第二分类相比正确地预测与第一类对应的输入数据的比例更大;

[0054] 基于第一分类和第二分类的输出来确定输入数据的分类预测。

[0055] 根据另一方面,提供了一种计算机实现的方法,包括:

[0056] 获得输入数据;

[0057] 将输入数据输入至分类模型,分类模型应用第一分类和第二分类,其中,第一分类使用第一预测阈值,并且第二分类使用第二预测阈值,其中,第一预测阈值高于第二预测阈值;

[0058] 基于第一分类和第二分类的输出来确定输入数据的分类预测。

[0059] 在示例中,第一分类包括应用第一分类器和使用第一预测阈值,并且第二分类包括应用第一分类器和使用第二预测阈值,其中,第一预测阈值高于第二预测阈值。

[0060] 在示例中,第一分类包括应用第一分类器,并且第二分类包括应用第二分类器。

[0061] 在示例中,输入数据包括组织图像,并且其中,确定分类包括确定与医疗诊断有关的信息。

[0062] 在示例中,基于第一分类和第二分类的输出来确定输入数据的分类预测包括:

[0063] 如果第一分类和第二分类的输出相同,则输出是第一分类或第二分类的输出的分类预测;以及

[0064] 确定不存在分类预测。

[0065] 根据另一方面,提供了一种计算机实现的分类方法,包括:

[0066] 获得输入数据;

[0067] 将输入数据输入至分类模型,分类模型应用第一分类,其中,第一分类包括应用第一分类器和使用与第一类对应的预测阈值;

[0068] 基于第一分类的输出来确定第一类的分类预测。

[0069] 在示例中,确定分类预测包括:

[0070] 如果第一分类的输出是第一类,则输出是第一类的分类预测;以及

[0071] 如果第一分类的输出不是第一类,则确定不存在分类预测。

[0072] 根据另一方面,提供了一种分类系统,该分类系统包括一个或更多个处理器,一个或更多个处理器被配置成执行上面的方法。

[0073] 根据另一方面,提供了一种分类系统,该分类系统包括根据上面的方法训练的分类模型。

[0074] 根据另一方面,提供了一种载体介质,该载体介质包括被配置成使计算机执行上面的方法的计算机可读代码。所述方法是计算机实现的方法。由于根据实施方式的一些方法可以通过软件实现,因此一些实施方法涵盖在任何合适的载体介质上提供给通用计算机的计算机代码。载体介质可以包括任何存储介质,例如软盘、CD ROM、磁装置或可编程存储装置,或任何瞬态介质,例如任何信号,如电信号、光信号或微波信号。载体介质可以包括非暂态计算机可读存储介质。

[0075] 分类器模型在广泛的领域中都具有应用。例如,分类器模型可以用于各种医疗应用中。分类器模型可以用于通过将例如生理测量值作为输入来提供医疗诊断。在另一示例中,分类器模型可以用于分析医学图像数据,例如用于预测输入组织图像是否包含癌细胞,或者用于预测特定生物标志物是否与组织图像相关联。

[0076] 这样的分类器模型可以使用训练数据集进行训练。例如,训练数据集可以包括大量示例输入以及标记的分类输出。然而,使用针对训练数据集已经建模的分类器模型的分方法在针对新输入数据部署时将具有一定程度的不准确度。

[0077] 例如,二元分类器模型针对每个数据点输出在0与1之间的得分,该得分基于预测阈值被转换成分类(正或负)。如果得分在阈值以上,则输出预测为正。如果得分在阈值以下,则输出预测为负。例如,阈值可以是0.5。例如,二元分类器模型可以用于预测全玻片图像是否包含癌细胞。当且仅当由分类器模型给出的得分大于或等于0.5时,输入的全玻片图像才会被预测为包含癌细胞。

[0078] 将分类器模型输出与预测阈值进行比较,以生成分类方法的输出(正或负)。分类方法的准确度可以由正确分类的数目占总分类数目的比例来测量。例如,在针对测试数据集中的每个输入已知正确的分类(正或负)的情况下,这可以针对该数据集进行评估。

[0079] 在一些使用实例例如医疗使用实例中,需要高的准确度。分类的准确度可以通过提高分类器模型性能来提高,其中分类器模型性能可以由AUC(下面将更详细地描述)来测量。这样的提高可以通过针对例如新的训练数据集进一步训练分类器模型来实现。然而,在一些情况下,例如在不能容易地获得另外的训练数据的情况下,不可能进一步提高分类器模型性能。

[0080] 图1示出了根据实施方式的分类方法的示意图。

[0081] 在S101中,获得输入数据。例如,输入数据可以是生理测量值或输入医学图像。

[0082] 在S102中,应用第一分类和第二分类。第一分类和第二分类各自将分类输出为正

类或负类。正类也称为第二类,并且负类也称为第一类。针对第一分类和第二分类,正类是相同的。例如,正类可能是医疗诊断,例如疾病的诊断。替选地,针对例如输入的组织图像,正类可以是癌细胞的检测或特定生物标志物的指示。第一分类被配置成与第二分类相比正确地预测更多的负实例。第二分类被配置成与第一分类相比正确地预测更多的正实例。第二分类被配置成比第一分类具有更高的真正率,其中真正率是该方法正确地预测为正的实例的比例。

[0083] 在S103中,然后基于第一分类和第二分类的输出确定分类预测。特别地,如果第一分类和第二分类的输出相同,则分类预测被输出为第一分类(或第二分类)的输出。如果第一分类和第二分类的输出不同,则不输出分类预测。然后,可以执行一些另外的步骤,例如人工分类。由于排除了预测更加不确定的情况,因此提高了分类预测的准确度。

[0084] 现在将描述分类的示例方法。

[0085] 第一示例

[0086] 存在分类方法的各种性能指标。分类方法的灵敏度是真正率(TPR),并且是该方法正确地预测为正的实例的比例。分类方法的特异度是真负率(TNR),并且是模型正确地预测为负的负实例的比例。方法的特异度也等于 $1 - \text{FPR}$,其中FPR是假正率。FPR是该方法错误地预测为正的负实例的比例。假负率(FNR)是模型错误地预测为负的正实例的比例,并且等于 $1 - \text{TPR}$ 。

[0087] 如上面所描述的,二元分类器模型针对每个数据点输出在0与1之间的得分,该得分基于预测阈值被转换成分类(正或负)。如果得分在阈值以上,则输出分类为正。如果预测得分在阈值以下,则输出分类为负。针对许多分类器模型,例如深度神经网络、支持向量机、随机森林等,在训练分类器模型之后,通过仅改变预测阈值,就可以调整分类方法的灵敏度和特异度。图2(a)是针对数据集的分类器模型输出的示意图。针对正类的分类器模型输出(概率值从0至1)以0.05的离散仓沿着x轴。针对数据集的每个分类器模型输出的频率对应于y轴。与真值为正(1)的实例对应的输出以较深阴影示出,而真值为负(0)的实例的输出以较浅阴影示出。

[0088] 较高预测阈值产生具有较高特异度的分类方法,该分类方法将能够正确地预测更多的负实例,但是将无法正确地预测更多的正实例。图2(a)中的线“特异度阈值”示出在其处数据集中所有真负实例被正确地预测的阈值。

[0089] 与较低预测阈值相结合的相同分类器模型产生具有较高灵敏度的分类方法,该分类方法将能够正确地预测更多的正实例,但是将无法正确地预测更多的负实例。图2(a)中的线“灵敏度阈值”示出在其处数据集中所有正实例被正确地预测的阈值。“特异度”阈值高于“灵敏度”阈值。

[0090] 使用分类器模型和较高阈值的分类方法被配置成:与使用分类器模型和较低阈值的分类方法相比,正确地预测更多的负实例。此外,基于分类器模型的灵敏度阈值的分类方法被配置成:与基于分类器模型的特异度阈值的分类方法相比,正确地预测更多的正实例。

[0091] 图2(b)示出了根据第一示例的方法。针对与组织的图像相对应的输入图像数据执行分类。在图3中示出了包括输入图像数据41的处理的示意图。然而,将理解,在其他示例中,可以使用其他类型的输入数据,例如包括以下其他类型中的一种或更多种:图像(例如非组织医学图像)、患者数据(例如年龄、性别、健康状况)、液体活检数据(例如DNA测序数

据)或血液测试数据(例如血细胞的数目)。

[0092] 在该示例中,输入图像被分类为其是否对应于特定生物标志物(ER生物标志物)。如果输入图像对应于ER生物标志物,则输入图像被分类成正类。然而,将理解,在其他示例中,可以执行其他类型的分类,例如以对图像是否对应于不同的生物标志物或癌细胞进行分类。

[0093] 根据第一示例的方法使用第一分类和第二分类。第一分类和第二分类使用相同的分类器。然而,第一分类的预测阈值高于第二分类的预测阈值。因此,第一分类被配置成与第二分类相比正确地预测更多的负实例。第二分类被配置成与第一分类相比正确地预测更多的正实例,换言之,第二分类被配置成与第一分类相比具有更高的真正率。

[0094] 在S201中,获得输入数据。在该示例中,输入数据是包括多个像素的输入图像数据41。如图3所示,该示例中的输入图像数据41是来自组织的全玻片图像的图像图块。在下面的描述中,像素数据是红绿蓝(尺寸为高度 $\times$ 宽度 $\times$ 3),然而像素数据可以替选地例如是灰度(尺寸为高度 $\times$ 宽度 $\times$ 1)。输入图像数据包括第一数目的像素,其中第一数目等于高度 $\times$ 宽度。可以使用捕获组织图像(也称为组织切片)的安装有显微镜的数字摄像装置获取图像数据。

[0095] 在本文描述的具体示例中,输入图像数据41包括利用苏木精和伊红染色剂染色的组织切片的图像的图块。图3(WSI)中示出了用利苏木精和伊红染色剂染色的组织切片的全玻片图像(WSI)的示例。在WSI上覆盖了网格,其中每个网格方块对应于图块。全玻片图像扫描仪扫描整个组织切片,产生利用苏木精和伊红染色剂染色的组织切片的图像,包括例如约60000像素高度 $\times$ 60000像素宽度。然后,全玻片图像被分割成连续的部分或图块。图像部分(图块)具有固定的高度和宽度。这些部分可以在图像内连续或交叠。例如,图像图块大小可以是512 $\times$ 512像素。全玻片图像被分割成图块,并且每个图块会被单独地用作S201中的输入。当然,可以使用其他图块大小。例如,可以使用与2的幂次方相对应的图块大小,例如:128 $\times$ 128像素、256 $\times$ 256像素、512 $\times$ 512像素或1024 $\times$ 1024像素。每个WSI图像可以具有不同的大小,并且因此可以根据输入WSI的大小从输入WSI提取不同数目的部分。原始图像数据可以作为输入,即原始像素值。替选地,可以对原始像素值执行一些预处理,例如以用于颜色归一化。最初可以执行细胞分割步骤,并且丢弃仅包含非癌症组织的图块。针对细胞分割可以使用预训练的模型。

[0096] 在S202中,将输入图像数据41(图像图块)作为分类模型的输入。分类模型包括第一分类器。在该示例中,第一分类器是第一卷积神经网络(CNN)40。在图4中示出示例第一卷积神经网络40。

[0097] 第一卷积神经网络40包括:第一部分46,该第一部分46包括至少一个卷积层;以及第二部分47,将一维向量作为输入的分类部分。第二部分47包括至少一个全连接层。第一CNN 40包括第一部分46,该第一部分46包括在各个深度层和视场层处应用于图像的卷积滤波器,接着是第二部分47,该第二部分47包括用于数据减少的全连接稠密层和/或池化层。滤波器权重是在训练阶段期间学习的可训练参数。当较低级滤波器检测诸如边缘和二进制大对象(blob)等粗结构时,较深级滤波器捕获如形状和纹理的更复杂的性质,以及最终顶层滤波器对于生物标志物的识别学习对感兴趣对象进行泛化。

[0098] 第一CNN 40是二元分类器。第一CNN 40用于确定输入图像图块是否与特定的分子

生物标志物相关联。图块被提交给第一CNN 40。每个像素的数据直接输入至第一CNN 40中。对于每个图块,CNN输出该图块被分配到正类,即该图块与生物标志物相关联的概率。

[0099] 图4中示出的第一CNN 40基于残差网络架构。残差神经网络包括一个或更多个跳过连接。为了简单起见,该图示出少量的层,然而,第一CNN 40可以包括例如超过100个层。

[0100] CNN中的第一层是卷积层,在图中标记为“卷积层1”。第一层中的每个滤波器都具有与输入数据的深度匹配的滤波器深度。例如,当输入数据为RGB时,第一层中的滤波器深度为3。为了简单起见,图4中示出的CNN的输入数据深度为1(即,灰度输入数据)。

[0101] 第一层的输出体积由多个因素确定。

[0102] 第一层的输出体积的深度对应于滤波器的数目。例如,在第一层中可以有32个滤波器,因此第一层的输出的深度为32。因此,后续层中的滤波器的深度也是32。

[0103] 输出体积的高和宽由输入的高度和宽度、滤波器的感受野大小(高和宽二者)以及滤波器步幅确定。当步幅为1时,那么滤波器每次滑动一个像素。当步幅为2时,那么滤波器每次滑动2个像素,从而产生较小的输出体积。边界处使用的任何零填充也将影响输出大小。

[0104] 每个滤波器沿着输入的宽度和高度移动。在每个位置处,进行输入值和权重值的逐元素乘法,然后对结果进行求和,以针对每个位置给出单个输出值。一个滤波器的输出值形成2D阵列。来自层中的所有滤波器的输出阵列沿深度维度堆叠,并且将得到的体积输入至下一层中。

[0105] 在每个卷积层之后为激活层。激活层应用逐元素激活函数,保持大小不改变。为了简单起见,图中未示出激活层。例如,模型可以包括一个或更多个(修正线性单元)层,该ReLU层应用逐元素激活函数。在每个卷积层之后还实现批量归一化层,其中在批量归一化层之后实现激活层。例如,模型包括包含以下的单元:卷积层、批量归一化层和激活层;或者包含以下的单元:第一卷积层、第一批量归一化层、第二卷积层、第二批量归一化层和激活层。

[0106] 第一CNN 40包括多个层,这些层的输出比输入具有更小的尺寸,例如高度和/或宽度比层的输入更小。以这种方式,通过大量层,高度和宽度减小,同时深度增大。第一CNN 40具有“编码器/解码器”结构,借此层首先(例如经由诸如步幅大小的滤波器超参数)减小高度和宽度,同时增大深度,并且然后(例如经由池化层和/或双线性上采样层)增大高度和宽度,同时减小深度。这在图4中示出,图4示出了层的输出大小。

[0107] 第一CNN 40还包括一个或更多个池化层,一个或更多个池化层被包括以改变空间大小。池化层用于增大输出的宽度和/或高度,并且减小深度。池化层可以是“平均池化”层。平均池化层包括具有空间范围和步幅的滤波器,该滤波器在输入上移动,取每个位置处的平均值。然而,可以使用除了平均值之外的函数,例如最大池化。还可以附加地或备选地包括上采样层,例如一个或更多个双线性上采样层,以便增大高度和/或宽度。

[0108] 第一CNN 40还包括至少一个跳过连接。实际上,该模型可以包括多个跳过连接,然而,为了简单起见,图4中示出了少量层和单个跳过连接。第二层“卷积层2”产生输出,该输出被称为输出m。第四层“卷积层4”产生输出o,该输出o与输出m具有相同尺寸。“卷积层5”的输入由第一层的输出m以及第四层的输出o产生。将来自较靠前的层的输出直接输入至较靠后的层就是“跳过连接”。该示例中的输出通过逐像素加法组合。在例如输出是不同大小的

情况下,备选地可以使用级联。使用一个或更多个跳过连接,将来自下游的信息直接馈送至上游。这在整个网络中保持高层次的全局和区域视觉特征。将来自较靠前的层的特征直接输入至较靠后的层中,跳过一个或更多个中间层提供了上下文。

[0109] 在最终卷积层之后包括展平层。展平层将来自最终卷积层的输出数据转换成1维向量 x ,以用于输入至下一层中。在该示例中,展平层之前的层形成CNN 46的第一部分。

[0110] 在展平层之后包括一个或更多个全连接层。最终全连接层输出与正类对应的值。在输出处应用激活函数,例如sigmoid函数,以给出概率值。激活函数将来自最终全连接层的输出作为输入,并且归一化为概率。因此,激活函数输出与正类对应的介于1与0之间的值。全连接层和激活函数形成第一CNN 40的第二部分47。

[0111] 对于每个图块,CNN 40输出该图块被分配到正类,即该图块与生物标志物相关联的概率。生物标志物是自然存在的分子、基因或特征,通过生物标志物可以识别特定的病理或生理过程、疾病、诊断、治疗或预后。在本文描述的具体示例中,生物标志物是癌症生物标志物,即自然存在的分子、基因或特征,通过癌症生物标志物可以识别特定类型的癌症或特别有效的癌症治疗。此外,在本文描述的示例中,生物标志物是分子生物标志物。在该示例中,特定生物标志物是ER生物标志物。然而,将在下面更详细地描述其他生物标志物的示例。

[0112] 在S203中,向第一分类器40的输出应用第一阈值。换言之,将从第一CNN 40输出的概率值与第一阈值进行比较。在图3中示出的示例中,从第一分类器输出的概率值为0.764。第一阈值是较高预测阈值T1。在该示例中,第一阈值T1为0.6。如果第一分类器的输出高于较高预测阈值,则第一分类输出为1,与正类对应。如果第一分类器的输出低于较高预测阈值,则第一分类输出为0。图6是示出第一预测阈值的示意图,该第一预测阈值在频率与分类器模型预测得分的曲线图上由线T1指示。

[0113] 在S203中,还向第一分类器的输出应用第二阈值。换言之,将从第一CNN 40输出的概率值单独地与第二阈值进行比较。第二阈值是较低预测阈值T2。在该示例中,第二阈值T2为0.45。如果第一分类器的输出高于较低预测阈值,则第二分类输出为1,与正类对应。如果第二分类器的输出低于较低预测阈值,则第二分类输出为0。第二预测阈值由图7中的线T2指示。

[0114] 第一分类的预测阈值T1高于第二分类的预测阈值T2。因此,第一分类被配置成比第二分类正确地预测更多的负实例。第二分类比第一分类正确地预测更多的正实例。

[0115] 第一分类的第一预测阈值T1高于第二分类的第二预测阈值T2。在使用系统之前执行的训练阶段期间确定阈值的值。下面将结合图8(a)描述训练阶段。

[0116] 在S204中,基于第一分类输出和第二分类输出确定最终分类预测。在该步骤中,如果第一分类输出1,并且第二分类输出1,则分类预测为1。如果第一分类输出0,并且第二分类输出0,则分类预测为0。如果第一分类和第二分类输出不同的分类,则不做出分类预测。可以执行一些另外的步骤来确定最终分类。例如,图像可以由专家手动分类。

[0117] 在该方法中,通过使用两个不同的预测阈值,可以使用单个利用单个训练过程训练的分类器执行两个不同的分类。第二分类M2使用较低预测阈值T2,并且因此提供具有高灵敏度的分类,产生更可靠的正预测。较低预测阈值在此还被称为负阈值。第一分类M1使用较高预测阈值T1,并且因此提供具有高特异度的分类,产生更可靠的负预测。较高预测阈值

在此还被称为正阈值。与通过使用第一分类器应用单个分类并且阈值为0.5相比,通过同时部署两个分类,可以实现分类预测中的更高程度的确定性。

[0118] 图8(a)是根据第一示例的训练分类系统的方法的示意图。该方法用于训练可以在根据上面描述的第一示例的分类方法中使用的系统。

[0119] 在训练方法中使用训练数据集和验证数据集,每个数据集包括多个输入图像。图像对应于在分类方法期间使用的预期类型的输入图像。在此处描述的示例中,输入图像是来自如上面已经描述的利用苏木精和伊红染色剂染色的组织切片的图像的图像图块。因此,训练数据集和验证数据集各自包括来自利用苏木精和伊红染色剂染色的组织切片的图像的图像图块。

[0120] 根据每个图像图块是否对应于模型要检测的特定生物标志物来标记每个图像图块。如上面所描述的,该示例中的特定生物标志物是ER生物标志物。如果每个图像图块对应于ER生物标志物,则用“1”标记该图像图块;而如果不对应,则用“0”标记该图像图块。

[0121] 为了生成标签,例如可以使用来自IHC染色处理的信息。对于一些数据集,例如在这些图像尚未作为元数据可用的情况下,专家可以查看IHC染色图像,并且确定目标图像的ER状态。然后在训练期间将ER状态用作用于对应H&E图像的真值标签。附加地或备选地,可以使用通过遗传学、转录组学和/或免疫学测定的方法对来自患者的人体样本的各种测试。这些测试以液体和/或固体形式在人体样本上进行,被称为活检,然后进行该过程以告知样本的分子状态。然后由专家——针对组织活检的病理学家、针对液体活检的血液学家、针对细胞学样本的细胞病理学家、针对遗传/转录组测定的遗传学家等——分析结果,以针对对应的H&E图像生成标签1或0。注释可以由受过训练的病理学家执行。

[0122] 为了准备数据集,执行与预期对输入图像执行的相同的图像预处理。因此,对于每个全玻片图像,获得多个图像部分(图块)。细胞分割可以用于从数据集中丢弃仅包含非癌组织的图块。

[0123] 可以针对全玻片图像的每个图块单独地执行图像注释。备选地,可以针对全玻片图像执行注释,并且然后利用对全玻片图像给出的注释对来自图像的每个图块进行标记。因此,在这种情况下,执行将每个图块与他们对应的切片的标签配对的步骤。

[0124] 因此,训练数据集和验证数据集各自包括多个图像图块,每个图块根据其是否对应于模型要检测的特定生物标志物而被标记。

[0125] 在S801中,以与上面结合推理(部署)阶段已经描述的方式相同的方式,训练数据集中的图像图块被提交给第一分类器,第一CNN 40,其生成图像被分配到正类的概率。

[0126] 第一CNN 40具有相关联的参数向量 θ_1 。参数包括第一CNN 40的第一部分中的所有卷积层的滤波器权重以及第一CNN 40的第二部分的权重。S801中的训练过程的目的是找到参数向量 θ_1' ,使得训练数据集中的图像的注释与第一分类器针对图像的输出之间的差异最小化。

[0127] 最优参数通过以下步骤来计算:首先赋随机值作为 θ_1 ;并且然后通过计算损失的梯度 $\frac{\partial D_1}{\partial \theta_1}$ 并且使用所计算的梯度更新 θ_1 来顺序地更新 θ_1 。 D_1 表示损失函数,该损失函数是“每图块”损失。该示例中使用交叉熵损失:

[0128] $-y_i \log(p_i) - (1-y_i) \log(1-p_i)$

[0129] 其中, y_i 是注释值 (1 或 0), 并且 p_i 是针对训练数据集中的示例 i 的分类器输出值。通过反向传播确定相对于分类器模型的可训练参数中的每一个的损失梯度。然后, 利用优化器函数, 使用梯度确定更新的参数。这一系列更新方法被称为梯度下降法 (GD), 通常迭代地定义为:

$$[0130] \quad \theta_1' = \theta_1 - \mu_1 \frac{\partial D_1}{\partial \theta_1}$$

[0131] 其中 μ_1 是学习率, 限定参数被更新的速度。更新可以基于批量平均值执行。例如, 可以使用 8 个图块的批量大小。计算 8 个图块的平均值损失, 并且基于平均损失更新模型参数。因此, 在更新参数之前, 分类器被应用于 8 个图块。

[0132] 可以使用 Adam 优化算法。然而, 可以根据每个策略针对使用实例的性能选择优化策略。例如, 可以选择以下优化方法之一:

[0133] • 随机梯度下降 (SGD)

[0134] • AdaDelta

[0135] • Adam

[0136] • AdaMax

[0137] • Nesterov Adam 优化器

[0138] • RMSProp

[0139] 可以在多个时期执行训练。时期的数目是分类器在训练期间应用于整个训练数据集的次数。在每个训练时期期间, 训练数据集中的训练示例可以以随机顺序输入至模型中, 使得针对每个时期以不同的顺序输入训练示例。

[0140] 可以在每个时期之后使用验证数据集评估分类器模型性能。在每个时期之后, 可以使用来自验证数据集的不同数据子集以用于评估。在每个训练时期之后, 可以确定基于 0.5 的阈值的分类器模型准确度。来自验证数据集的每个图像图块被提交给第一分类器, 该第一分类器生成图像被分配到正类的概率。然后将分类器模型输出与 0.5 的阈值进行比较, 以给出图像的分类。然后, 通过与图像的注释值 (标签) 进行比较, 确定分类是否正确。准确度被确定为正确分类的数目除以针对验证数据执行的分类的数目。准确度值是分类器模型性能的指示。一旦分类器模型已经实现期望性能, 则停止训练 (即不执行另外的时期), 并且分类器模型参数被固定。期望性能可以通过满足最小准确度 (例如, 0.8 或 0.9) 来确定, 或者通过当前时期与前一时期之间的准确度变化小于预定值来确定。代替监测准确度, 例如可以使用 AUC 或损失来以相同的方式执行监测。

[0141] 在 S802 中, 给定经训练的模型, 然后使用验证数据集确定预测阈值 T1 和 T2。在该步骤中, 假设验证集与测试集基本上相似。在 S802 中, 与上面结合推理 (部署) 阶段已经描述的方式相同的方式, 验证数据集中的图像图块被提交给第一分类器, 第一 CNN 40, 其生成图像被分配到正类的概率。然后将分类器模型预测得分, 即从分类器模型输出的概率值与图像的标签进行比较。在该示例中, 根据验证数据集结果, 将针对 (负类中的) 具有标签 0 的图像的最高概率值作为第一预测阈值 T1。将针对 (正类中的) 具有标签 1 的图像的最低概率值作为第二预测阈值 T2。因此, 如根据针对验证数据集运行第一分类器模型的结果所确定的, 第一预测阈值 T1 被设置为特异度阈值, 并且第二预测阈值 T2 被设置为灵敏度阈值。通过将第一预测阈值 T1 设置为特异度阈值并且将第二预测阈值 T2 设置为灵敏度阈值, 分类方法可

以针对正类和负类二者生成可靠预测。

[0142] 如上面结合图2(b)所描述的,可以使用另外的测试数据集来评估其中使用第一阈值T1应用第一分类和使用第二阈值T1应用第二分类的全分类方法的性能。例如,可以根据测试数据集确定全分类方法的准确度。

[0143] 在该示例中,第一预测阈值T1被设置为特异度阈值,并且第二预测阈值T2被设置为灵敏度阈值。然而,在替选示例中,第一预测阈值T1被设置为低于特异度阈值,并且第二预测阈值T2被设置为高于灵敏度阈值,同时仍确保第一预测阈值T1高于第二预测阈值T2。这增加了可以使用分类方法预测的实例的数目。例如,在医疗诊断系统中常见的是,如果收益超过不确定性或误差的成本,则可以引入可接受水平的这样的不确定性或误差。在灵敏度/特异度上下文中,可以通过将较低预测阈值T2(由M2使用)增大到略高于灵敏度阈值的值,并且将上预测阈值T1(由M1使用)减小到略低于特异度阈值来实现这一点。通过这样做,分类方法可以预测的实例的数目增加。

[0144] 给定具有相对较好性能的经训练的分类器模型,在如上面所描述的S802中确定预测阈值。在上面描述的假设验证集与测试集基本上相似的示例中,针对验证数据集,T1被设置为特异度阈值,并且T2被设置为灵敏度阈值,以便生成两个类的可靠预测。在阈值设置为灵敏度阈值的情况下,分类器模型针对验证数据集实现100%的PPV。PPV为正预测值,并且通过 $100 \times \text{TPR} / (\text{TPR} + \text{FPR})$ 给出。在阈值设置为特异度阈值的情况下,分类器模型针对验证数据集实现100%的NPV,其中,NPV是负预测值(NPV)并且通过 $100 \times \text{TNR} / (\text{FNR} + \text{TNR})$ 给出。

[0145] 图5是示出根据第一示例的经训练的分类系统的训练流水线和部署的示意图。如上面已经描述的,在S801中使用训练数据集训练第一分类器。然后针对验证数据集运行第一分类器,从而生成模型预测得分集,如图5所示出的。然后,在S802中基于模型预测得分集选择第一阈值T1和第二阈值T2。以自动方式确定第一阈值T1和第二阈值T2。例如,第一预测阈值T1被设置为特异度阈值,并且第二预测阈值T2被设置为灵敏度阈值。

[0146] 在S501中,然后针对测试数据集执行分类方法,以便评估分类方法的准确度。在S502中,确定针对仅对第一分类和第二分类一致的测试实例评估的测试数据集的分类准确度是否足以达到预期目的。例如,根据预期应用,0.90或0.95的准确度可能足够。如果准确度足够,那么如上面结合图2(b)所描述的,部署系统。如果不够,则系统被丢弃。

[0147] 在部署期间,在S201中接收输入图像,并且在S202中由第一分类器分析输入图像。在S203中输出第一分类和第二分类的预测。如果预测一致,则在S204中做出最终分类预测。如果不一致,则可以执行另外的步骤,例如手动测试。

[0148] 预测分类器模型的性能可以通过以下一项或两项来测量:

[0149] • 准确度:由分类器模型在特定预测阈值(例如0.5)的情况下做出的正确预测的数目/预测的数目;以及

[0150] • AUC:受试者操作特性曲线(ROC曲线)下的面积,该面积用作所有预测阈值的准确度的度量。

[0151] ROC曲线是针对不同阈值的真正率(TPR)与假正率(FPR)的曲线图。真正率也被称为灵敏度,并且是分类器模型利用阈值正确地预测为正的实例的比例。假正率是分类器模型利用阈值错误地预测为正的负实例的比例。使用测试数据集可以评估经训练的分类器模型的准确度(针对特定阈值,例如0.5)和AUC。例如,针对医疗使用实例,临床级医疗系统

可以使用具有0.94至0.98准确度和AUC的分类器模型。

[0152] 具有较低准确度和AUC的分类器模型可以用于以下应用:允许一定误差范围的预期用途,或者真预测的好处超过假预测的风险和影响,或者回路中存在确认预测的人。然而,针对许多使用实例,具有这样的分类器模型意味着装置在经济上不可行,例如,因为人必须确认预测。这可能阻碍在临床实践中采用该装置。例如,通过针对新的训练数据集进一步训练分类器模型,可以提高分类器模型性能。然而,在一些情况下,例如在不能容易地获得另外的训练数据的情况下,不可能进一步提高分类器模型性能。

[0153] 然而,分类方法的灵敏度和特异度可以通过选择与分类器模型一起使用的预测阈值来调整。图9示出了如通过从右到左遍历ROC曲线所观察到的从完全灵敏度(无假负)到完全特异度(无假正)的过渡的示意图。

[0154] 在上面描述的分类方法中,使用了如对验证数据集所评估的具有高或完全灵敏度(很少或没有假负)的第一分类以及具有高或完全特异度(很少或没有假正)的第二分类。因此,可以提高分类方法的准确度。

[0155] 在上面描述的示例中,使用验证数据集确定第一预测阈值和第二预测阈值。然而,在使用训练集、验证集和测试集的备选示例中,在S801中使用训练集执行训练,使用验证集验证训练的性能,并且然后使用测试集确定阈值。

[0156] 在结合图8(a)描述的训练方法中,训练数据集和验证数据集对应于同一群体。然而,在备选实施方式中,训练数据集和验证数据集可以对应于不同群体。例如,训练数据集可以对应于对其而言大量数据可用的普通群体。验证数据集可以对应于对其而言较小量数据可用的特定群体,并且该特定群体更紧密地对应于预期使用实例的群体。例如,训练数据集可以对应于来自一个国家或地区的人,而验证数据集可以对应于来自另一国家或地区的人。验证数据集可以对应于特定医院处的患者。验证数据集可以对应于具有特定特征或特性例如年龄范围的患者。以这种方式,使用大训练数据集训练第一分类器,换言之,使用大训练数据集更新第一分类器的参数。然后使用验证数据集来确定第一阈值和第二阈值。这可以被视为“校准”阶段,在“校准”阶段中,基于与较接近预期用于的群体的群体对应的验证数据集确定第一阈值和第二阈值。

[0157] 图8(b)是根据第一示例的训练分类系统的方法的示意图,其中验证数据集对应于与训练数据集不同的群体。该方法用于训练可以在根据上面描述的第一示例的分类方法中使用的系统。

[0158] 在训练方法中使用训练数据集和验证数据集,每个数据集包括多个输入图像。这些图像对应于在分类方法期间使用的预期类型的输入图像。在此处描述的示例中,输入图像是来自如上面已经描述的利用苏木精和伊红染色剂染色的组织切片的图像的图像图块。因此,训练数据集和验证数据集各自包括来自利用苏木精和伊红染色剂染色的组织切片的图像的图像图块。训练数据集包括对应于第一群体P1的图像,并且验证数据集包括对应于第二群体P2的图像。第二群体P2是目标群体,换言之,第二群体P2比第一群体P1更紧密地对应于预期使用实例的群体。例如,训练数据集可以包括来自所有年龄的患者的组织图像,而验证数据集包括来自特定年龄范围的患者组织图像。

[0159] 如上面所描述的,根据每个图像图块是否对应于模型要检测的特定生物标志物来标记每个图像图块。如上面所描述的,该示例中的特定生物标志物是ER生物标志物。如果每

个图像图块对应于ER生物标志物,则用“1”标记该图像图块;而如果不对应,则用“0”标记该图像图块。因此,训练数据集和验证数据集各自包括多个图像图块,每个图块根据其是否对应于要被检测的特定生物标志物而被标记。

[0160] 在S801中,将训练数据集中的图像图块提交给第一分类器,即第一CNN 40,其生成图像被分配到正类的概率。如上面所描述的,确定更新的参数向量 θ_1' ,使得训练数据集中的图像的注释与第一分类器针对图像的输出之间的差异最小化。使用训练数据集训练未训练的第一分类器,训练数据集对应于第一群体P1。S801的输出是针对第一群体P1训练的经训练的分类器。

[0161] 如上面所描述的,在每个训练时期之后,可以基于0.5的阈值确定分类器准确度。来自验证数据集的每个图像图块被提交给第一分类器,该第一分类器生成图像被分配到正类的概率。然后将分类器的输出与0.5的阈值进行比较,以给出图像的分类。附图中标记有“结果1”的第一张图表中示意性地示出了结果,该图示出了模型预测得分。所有在0.5以上的得分都被预测为正,而所有在0.5以下的得分都被预测为负。也在第一张图表中示出真分类,其中较深的阴影指示真值为正。

[0162] 通过与图像的注释值(标签)进行比较,可以确定分类是否正确。准确度被确定为正确分类的数目除以针对验证数据执行的分类的数目。准确度值是分类器模型性能的指示。一旦分类器模型已经实现期望性能,则停止训练(即不执行另外的时期),并且分类器参数被固定。

[0163] 在S802中,给定经训练的模型,然后使用验证数据集确定预测阈值T1和T2。验证数据集对应于第二群体P2,即目标群体。因此,可以将S802视为校准处理。将经训练的模型作为输入,并且确定预测阈值。

[0164] 如先前所描述的,在S802中,与上面结合推理(部署)阶段已经描述的方式相同的方式,验证数据集中的图像图块被提交给第一分类器,该第一分类器生成图像被分配到正类的概率。然后将分类器模型预测得分,即从分类器模型输出的概率值与图像的标签进行比较。在图表“结果2”中示出了分类器预测得分。在该示例中,根据验证数据集结果,将针对(负类中的)具有标签0的图像的最高概率值作为第一预测阈值T1。将针对(正类中的)具有标签1的图像的最低概率值作为第二预测阈值T2。因此,如根据针对验证数据集运行第一分类器模型的结果所确定的,第一预测阈值T1被设置为特异度阈值,并且第二预测阈值T2被设置为灵敏度阈值。通过将第一预测阈值T1设置为特异度阈值并且将第二预测阈值T2设置为灵敏度阈值,分类方法可以针对正类和负类二者生成可靠预测。

[0165] 如上面结合图2(b)所描述的,可以使用另外的测试数据集来评估其中使用第一阈值T1应用第一分类和使用第二阈值T1应用第二分类的全分类方法的性能。例如,可以根据测试数据集确定全分类方法的准确度。该示例中的测试数据集也对应于第二群体P2。

[0166] 在该示例中,第一预测阈值T1被设置为特异度阈值,并且第二预测阈值T2被设置为灵敏度阈值。然而,在备选示例中,第一预测阈值T1被设置为低于特异度阈值,并且第二预测阈值T2被设置为高于灵敏度阈值,同时仍确保第一预测阈值T1高于第二预测阈值T2。

[0167] S802中的校准步骤允许基于例如目标医院群体调整模型,换言之调整以适于群体转移。这通过以下来进行:确定预测阈值T1和T2,以适应于由群体变化引起的预测分数的分布变化。也可以包括调整模型的权重以适应于群体变化的中间步骤。例如,在使用第一训练

数据集执行训练过程之后,可以使用与第二群体P2对应的第二训练数据集执行另外的训练过程。在该另外的训练过程中,基于与目标群体对应的数据进一步调整第一分类器的参数。然后如上面所描述的,在S802中确定第一阈值和第二阈值。

[0168] 在上面描述的示例中,训练第一分类器,并且然后根据目标群体确定第一预测阈值和第二预测阈值。在校准阶段S802期间,将输入图像作为第一分类器的输入。在预测之前从最终层输出的值作为预测得分输出。因此生成目标群体预测得分分布。然后确定第一阈值和第二阈值。

[0169] 表1示出了根据第一示例的使用第一分类方法生成的结果,该第一分类方法将来自测试数据集的图像分类为对应于ER生物标志物。该方法使用如上面所描述的、具有第一阈值和第二阈值的分类器模型。表1还示出了根据比较示例的使用第一分类方法生成的结果,该第一分类方法将来自测试数据集的图像分类为对应于ER生物标志物。该方法使用如上面所描述的、具有单个阈值0.5的相同的分类器模型。

[0170] 表1还示出了根据第一示例的使用第二分类方法生成的结果,该第二分类方法将来自测试数据集的图像分类对应于PR生物标志物。该方法使用如上面所描述的、具有第一阈值和第二阈值的分类器模型。表1还示出了根据比较示例的使用第二分类方法生成的结果,该第二分类方法将来自测试数据集的图像分类为对应于PR生物标志物。该方法使用如上面所描述的、具有单个阈值0.5的相同的分类器模型。

[0171]

	比较示例		第一示例	
	ER 生物标志物	PR 生物标志物	ER 生物标志物	PR 生物标志物
准确度	0.82234957	0.768361582	0.870090634	0.826086957
F1	0.89084507	0.836653386	0.926243568	0.889655172
AUC	0.825141878	0.756788921	0.823986358	0.721239607
FPR	0.484848485	0.47008547	0.689655172	0.530612245
FNR	0.106007067	0.113924051	0.010989011	0.044444444
PPV	0.887719298	0.79245283	0.870967742	0.832258065

[0172]

NPV	0.53125	0.696629213	0.857142857	0.793103448
灵敏度	0.893992933	0.886075949	0.989010989	0.955555556
特异度	0.515151515	0.52991453	0.310344828	0.469387755
PLR	1.843860424	1.884925201	1.434065934	1.800854701
NLR	0.205778424	0.214985708	0.035409035	0.09468599

[0173] 表1

[0174] 在上面的表1中,PLR是正似然比,而NLR是负似然比。

[0175] 第一示例的修改

[0176] 图10(a)是根据第一示例的修改的训练分类系统的方法的示意图,其中较低预测阈值T2可以设置为高于灵敏度阈值的值,并且上预测阈值T1可以设置为低于特异度阈值的值。通过这样做,可以增加模型可以预测的实例的数目。

[0177] 在S1001中,以与上面结合S801描述的方式相同的方式使用训练数据集训练第一分类器40。

[0178] 在S1002中,以与上面结合S802描述的方式相同的方式使用验证数据集确定灵敏度阈值和特异度阈值。

[0179] 在S1003中,使用优化处理确定第一阈值和第二阈值。在根据第一示例的修改的方法中,代替将第一阈值和第二阈值分别设置为特异度阈值和灵敏度阈值,执行优化处理来确定第一阈值和第二阈值。

[0180] 现在将结合图10(b)描述可以在S1003中执行的示例优化处理。

[0181] 在S901中,针对第一阈值获得第一可能值集,并且针对第二阈值获得第二可能值集。第一集和第二集中的所有值都大于或等于灵敏度阈值,并且小于或等于特异度阈值。第一可能值集通过在灵敏度阈值与特异度阈值之间取固定数目的均匀间隔的值来获得,第一可能值集包括特异度阈值但是不包括灵敏度阈值。第二可能值集通过在灵敏度阈值与特异度阈值之间取固定数目的均匀间隔的值来获得,第二可能值集包括灵敏度阈值但是不包括特异度阈值。

[0182] 在该示例中,特异度阈值为0.65,并且灵敏度阈值为0.4。将以0.025为增量的一组均匀间隔的值作为第一值集和第二值集。第一值集包括值{0.425,0.450,0.475,0.500,0.525,0.550,0.575,0.600,0.625,0.650}。第二值集包括值{0.400,0.425,0.450,0.475,0.500,0.525,0.550,0.575,0.600,0.625}。

[0183] 在S902中,针对可能的第一阈值和可能的第二阈值的多个组合中的每个组合计算准确度值。根据测试数据集确定准确度。在该步骤中,将来自第一集中的每个值与来自第二集中的每个较低值进行组合。来自测试数据集的输入通过第一分类器运行,并且将来自第一分类器的输出概率值与组合中的第一阈值和第二阈值进行比较,以给出分类预测。然后将组合的准确度确定为正确分类预测的数目除以分类预测的总数目。针对多个组合中的每个组合重复上述步骤。

[0184] 针对图10(c)中的当前示例,示出了该步骤的结果。沿着网格顶部示出了第一集{0.425,0.450,0.475,0.500,0.525,0.550,0.575,0.600,0.625,0.650}中的每个值。顺着网格的左手侧示出了第二集{0.400,0.425,0.450,0.475,0.500,0.525,0.550,0.575,0.600,0.625}中的每个值。网格中的其余单元格示出了准确度值(针对与可能组合对应的单元)或者为空(针对与来自第二集的值低于来自第一集的值组合对应的单元)。来自第一集的每个值都与来自第二集的每个较低值进行组合,得到以下组合:具有对应准确度值的(0.425,0.400)、(0.450,0.400)、(0.450,0.425)、……(0.650,0.600)、(0.652,0.600)。针对每个组合,第一阈值都高于第二阈值。图10(c)示出了针对组合中的每个组合计算的准确度值。

[0185] 在S903中,检索针对预期应用产生准确度得分大于或等于期望准确度的第一阈值和第二阈值的所有组合。期望准确度是预定的准确度值。在该示例中,期望准确度为0.9。在该示例中,组合(0.575,0.400)、(0.600,0.400)、(0.600,0.425)、(0.625,0.400)、(0.625,0.425)、(0.625,0.450)、(0.625,0.475)、(0.650,0.400)、(0.650,0.425)、(0.650,0.450)、(0.650,0.475)产生大于0.9的准确度。因此在S903中检索到这些组合。

[0186] 在S904中,从检索到的组合中选择一个组合。基于覆盖范围选择组合。例如,选择所检索到的组合中的针对来自测试数据集的最大数目的结果预测分类的组合。备选地,选择第一阈值与第二阈值之间的差最大的组合。在该示例中,组合(0.625,0.475)被选择为具有最大覆盖范围。因此,在S1003中,第一阈值被设置为0.625,并且第二阈值被设置为0.475。

[0187] 第二示例

[0188] 图11(a)示出了根据第二示例的分类方法。在结合图11(a)描述的方法中,针对结合第一示例描述的相同输入图像数据执行分类方法。图11(b)示出了包括输入数据的该方法的示意图。

[0189] 根据第二示例的方法也使用第一分类和第二分类。然而,第一分类和第二分类使用不同的分类器。第一分类法使用被配置成比用于第二分类的第二分类器正确地预测更多负实例的第一分类器。第二分类器被配置成比第一分类器具有更高的真正率。在第二实施方式中,训练了两个单独的分类器,其中一个具有高特异度和低灵敏度,而另一个具有高灵敏度和低特异度。

[0190] 在S1101中,以与上面结合S201描述的方式相同的方式获得输入数据。

[0191] 在S1102中,将输入图像数据41(图像图块)作为分类模型的输入。分类模型包括第一分类器和第二分类器。因此,将输入图像数据41(图像图块)用作第一分类器的输入,并且还单独地作为第二分类器的输入。在该示例中,第一分类器是如上面结合图4已经描述的卷积神经网络(CNN)40。第二分类器也是如上面结合图4已经描述的卷积神经网络(CNN)40。在

该示例中,第一分类器和第二分类器具有相同的模型架构,但是使用不同的训练数据集进行训练,并且因此可以具有不同的参数值。如上面结合图4已经描述的分类器的结构是相同的。不过,在备选实施方式中,第一分类器和第二分类器具有不同的模型架构。例如,第一分类器和第二分类器可以具有不同的层数目、层中不同的过滤器数目或者不同的跳过连接。

[0192] 在S1103中,向第一分类器40a的输出应用第一阈值。换言之,将从第一CNN 40a输出的概率值与第一阈值进行比较。在图11(b)中示出的示例中,从第一分类器40a输出的概率值为0.764。在该示例中,与第一分类器40a一起使用的第一预测阈值是根据验证数据集确定的第一分类器40a的特异度阈值。如果第一分类器的输出高于第一预测阈值,则第一分类输出为1。如果第一分类器的输出低于第一预测阈值,则第一分类输出为0。

[0193] 在S1103中,向第二分类器40b的输出应用第二阈值。换言之,将从第二CNN 40b输出的概率值单独地与第二阈值进行比较。在该示例中,与第二分类器40b一起使用的第二预测阈值是根据验证数据集确定的第二分类器40b的灵敏度阈值。如果第二分类器40b的输出高于第二预测阈值,则第二分类输出为1。如果第二分类器的输出低于第二预测阈值,则第二分类输出为0。

[0194] 在该示例中,与第一分类器40a一起使用的第一预测阈值是根据验证数据集确定的第一分类器40a的特异度阈值,并且与第二分类器40b一起使用的第二预测阈值是根据验证数据集确定的第二分类器40b的灵敏度阈值。将在下面结合图12描述如何确定这些阈值。然而,在备选示例中,第一阈值和第二阈值被设置为相同的预定值。例如,第一阈值和第二阈值可以被设置为0.5。

[0195] 第一分类被配置成比第二分类正确地预测更多的负实例。第二分类被配置成比第一分类正确地预测更多的正实例。用于第一分类的被配置成正确地预测更多的负实例的第一分类器40a在部署之前执行的训练阶段期间获得。用于第二分类的被配置成正确地预测更多的正实例的第二分类器40b也在训练阶段期间获得。第二分类器40b被选择成比第一分类器具有更高的真正率。

[0196] 在S1104中,基于第一分类输出和第二分类输出确定分类预测。在该步骤中,如果第一分类输出1,并且第二分类输出1,则分类预测为1。如果第一分类输出0,并且第二分类输出0,则分类预测为0。如果第一分类和第二分类输出不同的分类,则不做出分类预测。然后可以执行一些另外的步骤来确定最终分类。例如,图像可以由专家手动分类。

[0197] 图12(a)是根据第二示例的训练分类系统的方法的示意图。该方法用于训练可以在根据上面描述的第二示例的分类方法中使用的分类系统。

[0198] 使用训练数据集和验证数据集,每个数据集包括多个图像。这些图像对应于预期类型的输入图像。如先前所描述的,训练数据集和验证数据集可以对应于相同的群体,或者训练数据集可以对应于第一群体,并且验证数据集可以对应于第二群体(目标群体)。

[0199] 在此处描述的示例中,输入图像是来自利用苏木精和伊红染色剂染色的组织切片的图像的图像图块。因此,训练数据集和验证数据集各自包括来自利用苏木精和伊红染色剂染色的组织切片的图像的图像图块。如结合第一示例所描述的,每个图像图块也根据其是否对应于模型要检测的特定生物标志物而被标记。

[0200] 在此处描述的示例中,训练数据集被分割成五个单独的训练数据子集。训练数据集中的每个图像都被随机分配到五个单独的训练数据子集中的一个。因此,每个训练数据

子集都包括不同的图像。每个子集具有相同数目或基本上相同数目的图像。

[0201] 在S1201中,对A、B、C、D和E五个单独的分类器进行训练。每个分类器具有相同的模型架构,但是使用不同的训练数据子集。

[0202] 来自第一训练数据子集的图像图块被提交给分类器A,该分类器A生成被分配到正类的概率。分类器A具有相关联的参数向量 θ_A 。最优参数通过以下来计算:将随机值分配为 θ_A ,并且然后通过计算损失的梯度 $\frac{\partial DA}{\partial \theta_A}$ 并且使用所计算的梯度更新 θ_A 来顺序地更新 θ_A 。DA表示损失函数,该损失函数是“每图块”损失。如先前所描述的,使用交叉熵损失。通过反向传播确定相对于第一分类器的可训练参数中的每一个的损失的梯度。然后,如先前已经描述的,利用优化器函数,使用梯度确定更新的参数。

[0203] 来自第二训练数据子集的图像图块被提交给分类器B,该分类器B生成被分配到正类的概率。分类器B具有相关联的参数向量 θ_B 。最优参数通过以下来计算:将随机值分配为 θ_B ,并且然后通过计算损失的梯度 $\frac{\partial DB}{\partial \theta_B}$ 并且使用所计算的梯度更新 θ_B 来顺序地更新 θ_B 。在该示例中,使用相同形式的损失函数训练每个分类器。通过反向传播确定相对于模型的可训练参数中的每一个的损失的梯度。然后利用如先前所描述的优化器函数,使用梯度来确定更新的参数。在该示例中,对于每个分类器使用相同的学习率 μ 。

[0204] 第三训练数据子集、第四训练数据子集和第五训练数据子集分别用于以相同的方式训练分类器C、D和E。因此,在该示例中,五个分类器之间的唯一差异是在S1201中用于训练过程的训练数据。

[0205] 对于每个分类器,使用如先前所描述的来自验证数据集的数据对模型性能进行评估。一旦训练已经实现期望性能,对于每个分类器而言,模型参数是固定的。

[0206] 在S1202中,来自验证数据集的相同数据被输入至五个分类器中的每一个。然后将预测得分,即从分类器输出的概率值与图像标签进行比较。然后针对每个分类器计算特异度阈值和灵敏度阈值。对于每个分类器,验证数据集结果用于确定针对(负类中的)具有标签0的图像的最高输出概率值——这就是特异度阈值。对于每个分类器,针对(正类中的)具有标签1的图像的最低输出概率值被确定为灵敏度阈值。与最高特异度阈值对应的分类器被选择为第一分类器,该第一分类器将能够比其他分类器正确地预测更多的负实例。具有最低灵敏度阈值的分类器被选择为第二分类器,该第二分类器将能够比其他分类器正确地预测更多的正实例。与第一分类器一起使用的第一预测阈值被设置为所选择的分类器的特异度阈值。与第二分类器一起使用的第二预测阈值被设置为所选择的分类器的灵敏度阈值。在该第二示例中,选择了两个模型,第一模型用于预测负实例,并且第二模型用于预测正实例。然后针对第一模型选择特异度阈值,并且针对第二模型选择灵敏度阈值。该方法可以用于其中一些模型仅非常擅长预测负实例而不擅长预测正实例而其他模型非常擅长预测正实例而不擅长预测负实例的应用。

[0207] 然后,可以使用测试数据集评估其中如结合图11(a)所描述的使用第一分类器应用第一分类并且使用第二分类器应用第二分类的全分类的性能。

[0208] 在上面的示例中,训练了五个分类器,其中与最高特异度阈值对应的分类器被选择为第一分类器,并且具有最低灵敏度阈值的分类器被选择为第二分类器。然而,在备选示例中,仅训练了两个分类器,其中与较高特异度阈值对应的分类器被选为第一分类器,并且

另一分类器被选择为第二分类器。在备选示例中,训练了多于两个分类器,其中与最高特异度阈值对应的分类器被选择为第一分类器,并且具有最低灵敏度阈值的分类器被选择为第二分类器。

[0209] 在上面的示例中,与最高特异度阈值对应的分类器被选择为第一分类器,并且具有最低灵敏度阈值的分类器被选择为第二分类器。然而,备选地,使用0.5的阈值的与最高特异度对应的分类器被选择为第一分类器,并且使用0.5的阈值的具有最高灵敏度的分类器被选择为第二分类器。然后将第一预测阈值和第二预测阈值二者都设置为0.5。

[0210] 在上面的示例中,多个分类器中的每个分类器都利用训练数据集的单独子集进行训练,其中每个图像被随机分配到子集中的一个。然而,在备选示例中,可以基于图像的一个或更多个性质来分配图像。例如,子集中的一个或更多个可以比其他子集具有更多正示例。例如,可以分配图像,使得子集中的两个具有80%的正示例,子集中的一个具有50%的正示例,以及子集中的两个具有20%的正示例。在上面的示例中,子集具有基本上相同数目的图像,然而备选地,子集中的一些可以比其他子集具有更多的图像。

[0211] 在上面的示例中,多个分类器中的每个分类器都具有相同的模型架构。然而,在备选示例中,分类器中的一些或全部可以具有不同的模型架构,例如,分类器可以具有不同数目的层、层中不同数目的过滤器或不同的跳过连接。

[0212] 在上面的示例中,多个分类器中的每一个都使用相同的超参数进行训练。然而,在备选示例中,可以使用不同的超参数训练分类器中的一些或全部,例如学习率可以不同。

[0213] 图12(b)是示出根据第二示例的训练流水线和经训练的分类系统的部署的示意图。如上面已经描述的,在S1201中使用训练数据集训练两个分类器。然后针对验证数据集运行分类器,从而针对每个分类器生成模型预测得分集,如图12(b)所示出的。然后,基于模型预测得分集,选择第一分类器和第二分类器。以自动方式确定第一分类器和第二分类器。例如,如基于验证数据集结果所确定的,第一分类器被选择为与最高特异度阈值对应的分类器,并且第二分类器被选择为具有最低灵敏度阈值的分类器。

[0214] 在S501中,针对测试数据集执行全分类方法,以便评估该方法的准确度。在S502中,确定针对仅对声称确定(即,第一分类和第二分类是否一致)的测试实例评估的测试数据集的准确度是否足以达到预期目的。如果准确度足够,那么如上面结合图11(a)所描述的,部署系统。如果不够,则系统被丢弃。

[0215] 在部署期间,在S1101中接收输入图像,并在S1102中由第一分类器和第二分类器分析输入图像。在S1103中输出第一分类和第二分类的预测。如果预测一致,则在S1104中做出最终分类预测。如果不一致,则可以执行另外的步骤,例如手动测试。

[0216] 第二示例的修改

[0217] 图12(c)是根据第二示例的修改的训练分类器的方法的示意图。

[0218] 在S1401中,与上面结合S1201描述的方式相同的方式训练多个分类器。

[0219] 在S1402中,与上面结合S1202描述的方式相同的方式选择第一分类器和第二分类器。如结合S1202所描述的,作为该步骤的一部分,使用验证数据集确定第二分类器的灵敏度阈值和第一分类器的特异度阈值。

[0220] 在上面描述的第二示例中,与第一分类器40a一起使用的的第一预测阈值被设置为根据验证数据集确定的第一分类器40a的特异度阈值,并且与第二分类器40b一起使用的第

二预测阈值被设置为根据验证数据集确定的第二分类器40b的灵敏度阈值。然而,在根据第二示例的修改的方法中,执行优化处理以确定第一阈值和第二阈值。

[0221] 在S1403中,使用优化处理确定第一阈值和第二阈值。在该步骤中,可以执行上面结合图10(b)所描述的示例优化处理。

[0222] 针对第一阈值获得第一可能值集,并且针对第二阈值获得第二可能值集。在该示例中,第一集中的所有值都大于或等于第一分类器的灵敏度阈值,并且小于或等于第一分类器的特异度阈值,而第二集中的所有值都大于或等于第二分类器的灵敏度阈值,并且小于或等于第二分类器的特异度阈值。例如,第一可能值集可以通过从第一分类器的灵敏度阈值到第一分类器的特异度阈值取固定数目的均匀间隔的值来获得,第一可能值集包括特异度阈值和灵敏度阈值。第二可能值集可以通过从第二分类器的灵敏度阈值到第二分类器的特异度阈值取固定数目的均匀间隔的值来获得,第二可能值集包括灵敏度阈值和特异度阈值。

[0223] 然后,针对可能的第一阈值和可能的第二阈值的多个组合中的每个组合确定准确度。在该步骤中,来自第一集的每个值与来自第二集的每个值进行组合。来自测试数据集的输入通过第一分类器和第二分类器运行。第一分类器模型输出与组合中的第一阈值进行比较,并且第二分类器模型输出与组合中的第二阈值进行比较,以给出分类预测。然后针对组合中的每一个确定准确度。

[0224] 检索产生准确度得分大于或等于针对应用的期望准确度的第一阈值和第二阈值的所有组合,并且然后基于覆盖范围从所检索到的组合中选择组合。例如,选择所检索到的组合中的根据测试数据集分类出最多结果的组合。

[0225] 第三示例

[0226] 在第一示例中,在每个图块的基础上应用分类方法。从全玻片图像中取图像图块,并且被作为第一分类器的输入。应用使用第一阈值的第一分类和使用第二阈值的第二分类。然后确定关于图像图块是否对应于ER生物标志物的分类预测。

[0227] 图13是根据第三示例的分类方法的示意图,其中在每个图像的基础上应用分类方法。换言之,获得关于全玻片图像是否对应于ER生物标志物的分类预测。

[0228] 在S1501中,获得输入数据。该示例中的输入图像数据是来自如先前所描述的组织的全玻片图像的图像图块。在该示例中,如结合图13所示出的,从全玻片图像获得多个图像图块。细胞分割可以用于丢弃仅包括非癌症组织的图块。

[0229] 在S1502中,将每个输入图像图块单独地作为分类模型的输入,该分类模型包括第一分类器,该第一分类器在该示例中是如结合第一示例所描述的第一CNN 40。第一分类器针对每个图像图块输出图像图块与ER生物标志物相关联的概率。

[0230] 然后执行聚合步骤。在该示例中,聚合步骤包括应用池化算子。识别与最高概率值对应的图像图块。该图像图块将被称为“排名靠前”的图像图块。然后,在分类方法的后续步骤中,使用该“排名靠前”的图像图块对图像进行分类。

[0231] 在S1503中,将第一阈值应用于第一分类器40针对排名靠前的图像图块的输出。换言之,将从第一CNN 40针对排名靠前的图像图块输出的概率值与第一阈值进行比较,以给出第一分类的输出。将第二阈值应用于第一分类器40针对排名靠前的图像图块的输出。换言之,将从第一分类器40针对排名靠前的图像图块输出的概率值单独地与第二阈值进行比

较,以给出第二分类。第一阈值和第二阈值是在训练阶段期间确定的,如先前结合第一示例所描述的。

[0232] 在S1504中,基于第一分类输出和第二分类输出确定全玻片图像的最终分类预测。在该步骤中,如果第一分类输出1,并且第二分类输出1,则分类预测为1。如果第一分类输出0,并且第二分类输出0,则分类预测为0。如果第一分类和第二分类输出不同的分类,则不做出分类预测。可以执行一些另外的步骤来确定最终分类。例如,图像可以由专家手动分类。

[0233] 在第三示例中,在部署期间,第一CNN分类器40在每个图块的基础上返回概率,并且这些单独的得分通过最大算子进行聚合。替选地,可以使用其他不可训练的聚合函数,例如求平均值。第一分类器40以与针对第一示例相同的方式进行训练。将针对全玻片图像的分类预测作为针对单个“代表性”图块的分类预测。

[0234] 在替选示例中,使用可训练的聚合函数。通过将来自多个图块的数据输入至聚合模型中,确定输入的全玻片图像是否与生物标志物相关联的指示。聚合模型可以包括递归神经网络,例如长短期记忆(LSTM)网络。第一分类器40的输出可以用于基于从第一分类器40输出的概率选择减少的图像图块集,例如前k个图块,其中k是大于2的整数。将减少的图像图块集作为分类模型的输入。首先使用第一经训练的模型生成与图像图块中的每一个对应的特征向量。然后将这些特征向量依次作为第二训练的模型的输入,该第二经训练的模型可以包括例如LSTM。LSTM的与序列中的最终输入对应的输出被输入至全连接神经网络层,该全连接神经网络层输出输入图像与ER生物标志物相关联(与正类对应)的概率。然后将该输出值与第一阈值进行比较以给出第一分类的输出,并且与第二阈值进行比较以给出第二分类的输出。然后基于第一分类和第二分类的输出确定针对全玻片图像的分类预测。在训练期间,来自训练数据集的图像图块将以与在推理期间相同的方式作为分类模型的输入。将输出与针对全玻片图像的标签进行比较,以便使用例如如先前已经描述的交叉熵损失函数和优化器函数来训练第一模型和第二模型。在替选示例中,分类模型可以包括应用包括具有第一阈值的第一经训练的模型和第二经训练的模型的第一分类,以及包括具有第二阈值的不同的第一经训练的模型和不同的第二经训练的模型的第二分类。

[0235] 第四示例

[0236] 在上面描述的示例中,针对所有的类获得了提高的准确度,但是预测覆盖范围,即分类模型可以预测的实例的数目减小。然而,在一些应用中,分类方法仅用于预测类中的一些类中,例如仅预测负类或仅预测正类。例如,分类方法可以用于对图像是否对应于HER2生物标志物进行分类,其中模型仅用于预测负类。

[0237] 在这样的示例中,分类方法可以使用单个分类器和单个预测阈值。针对某些类获得提高的准确度。例如,分类器仅用于预测第一类,其中针对第一类获得提高的准确度。在方法中使用了与第一类对应的阈值。如果分类预测第一类,则做出第一类的分类预测。如果分类预测不预测第一类,则不做出分类预测。

[0238] 如先前已经描述的,较高预测阈值产生具有较高特异度的分类方法,该分类方法将能够正确地预测更多的负实例,但是将无法正确地预测更多的正实例。例如,图2(a)中的线“特异度阈值”示出在其处数据集中所有真负实例被正确地预测的阈值。与较低预测阈值相结合的相同分类器产生具有较高灵敏度的分类方法,该分类方法将能够正确地预测更多的正实例,但是将无法正确地预测更多的负实例。图2(a)中的线“灵敏度阈值”示出在其处

数据集中所有正实例被正确地预测的阈值。“特异度”阈值高于“灵敏度”阈值。

[0239] 图16示出了根据第四示例的方法。如结合第一示例所描述的,针对与组织图像对应的输入图像数据执行分类。如果输入图像不对应于HER2生物标志物,则将输入图像分类至负类中。

[0240] 根据第四示例的方法使用第一分类。

[0241] 在S1601中,以与S201中结合第一示例描述的方式相同的方式获得输入数据。

[0242] 在S1602中,将输入数据作为分类模型的输入。分类模型应用第一分类器。如第一示例中结合S202所描述的,第一分类器可以是CNN。

[0243] 在S1603中,向第一分类器的输出应用阈值。换言之,将从第一分类器输出的概率值与阈值进行比较。在该示例中阈值对应于用于负类的阈值。例如,阈值可以对应于如使用数据集所确定的灵敏度阈值。

[0244] 在S1604中,基于来自S1603的第一分类输出确定针对负类的最终分类预测。在该步骤中,如果第一分类输出0,则分类预测为0。如果第一分类输出1,那么不做出分类预测。可以执行一些另外的步骤来确定最终分类。例如,图像可以由专家手动分类。

[0245] 在该示例中,第一分类用于预测负类。在替选的示例中,第一分类用于预测正类。在这种情况下,使用较高预测阈值,使得阈值与针对正类的阈值对应。例如,阈值可以对应于如使用数据集确定的特异度阈值。

[0246] 图17是根据第四示例的训练分类系统的方法的示意图。该方法用于训练可以在根据上面描述的第一示例的分类方法中使用的系统。

[0247] 在训练方法中使用训练数据集和验证数据集,每个数据集包括多个输入数据。如先前所描述的,输入数据对应于在分类方法期间使用的预期类型的输入数据。如先前已经描述的,训练数据集可以对应于第一群体,并且验证数据集可以对应于目标群体。在该示例中,如结合第一示例已经描述的,输入数据包括根据其是否对应于模型要检测的特定生物标志物而被标记的组织图像。如上面所描述的,该示例中的特定生物标志物是HER2生物标志物。如果每个图像图块对应于生物标志物,则用1标记该图像图块;而如果不对应,则用0标记该图像图块。

[0248] 在S1701中,将训练数据集中的图像图块提交给第一分类器,即第一CNN 40,该第一分类器生成图像被分配到正类的概率。如上面所描述的,确定更新的参数向量 θ_1' ,使得训练数据集中的图像的注释与第一分类器针对图像的输出之间的差异最小化。

[0249] 在S1702中,给定经训练的模型,然后使用验证数据集确定预测阈值。以与上面结合推理(部署)阶段已经描述的方式相同的方式,验证数据集中的图像图块被提交给第一分类器,第一CNN 40,该第一分类器生成图像被分配到正类的概率。然后将分类器模型预测得分,即从分类器模型输出的概率值与图像的标签进行比较。在该示例中,根据验证数据集结果,将针对(正类中的)具有标签1的图像的最低概率值作为预测阈值。因此,如根据针对验证数据集运行第一分类器模型的结果所确定的,预测阈值被设置为灵敏度阈值。通过将预测阈值设置为灵敏度阈值,分类方法可以生成负类的准确预测,换言之,假负率,即模型错误地预测为负的正实例的比例降低。

[0250] 在上面描述的示例中,预测阈值被设置为如根据验证数据集确定的灵敏度阈值。然而,在一些示例中,预测阈值被设置为高于灵敏度阈值的值。例如,预测阈值可以使用优

化处理来确定。针对阈值获得可能值集。在该示例中,如先前结合第一示例的修改已经描述的,集中的所有值都大于或等于第一分类器的灵敏度阈值,并且小于或等于第一分类器的特异度阈值。

[0251] 然后针对每个可能的阈值确定得分。例如,得分可以是FNR或NPV。来自测试数据集的输入通过第一分类器运行。将第一分类器模型输出与阈值进行比较,以给出分类预测。然后,例如通过确定预测集的FNR,针对阈值确定得分。针对每个可能的阈值重复以上步骤。

[0252] 检索产生得分大于或等于针对应用的期望得分的所有阈值,并且基于覆盖范围从所检索的阈值中选择阈值。例如,根据测试数据集,从所检索的阈值中选择分类出最多结果的阈值,或者选择最高阈值。

[0253] 在第一分类用于预测正类的示例中,预测阈值可以设置为如根据验证数据集确定的特异度阈值,以便降低假正率。替代地,可以通过优化处理以如上面所描述的类似方式设置阈值。

[0254] 在第四示例中,分类方法使用单个分类器,并且可以根据目标群体,例如使用与目标群体对应的验证数据集来调整阈值。示例使用具有一个阈值的一个分类器,其中分类方法旨在例如仅预测负类或仅预测正类。在一些示例中,针对训练集训练分类器,针对验证集验证训练的性能,以及针对测试集设置阈值。

[0255] 表2示出了根据第四示例的使用第一分类方法生成的结果,该第一分类方法将来自测试数据集的图像分类为不与HER2生物标志物对应。该方法使用如上面所描述的、具有与负类对应的单个阈值的分类器模型。表2还示出了根据比较示例的使用第一分类方法生成的结果,该第一分类方法将来自测试数据集的图像分类为与HER2生物标志物对应。该方法使用如上面所描述的、具有单个阈值0.5的相同的分类器模型。

[0256]

	比较示例	第四示例
准确度	0.647482014	0.867521368
F1	0.319444444	0.114285714
AUC	0.612370114	0.558557127
FPR	0.331914894	0.009852217
FNR	0.465116279	0.935483871
PPV	0.227722772	0.5
NPV	0.88700565	0.873913043
灵敏度	0.534883721	0.064516129
特异度	0.668085106	0.990147783

[0257]

PLR	1.611508646	6.548387097
NLR	0.696193157	0.944792168

[0258] 表2

[0259] 通过了解通常癌症和/或患者的特异性分子谱,可以了解针对癌症进行的各种外科手术,包括激素疗法、免疫疗法或靶向药物治疗等。各种医学相关的生物标志物,包括诊断标志物、治疗标志物或预后标志物中的任一种,包括突变状态、受体状态或拷贝数变异等,都可以被识别以确定患者对某些疗法的反应如何。突变状态、受体状态或拷贝数变异都

是分子生物标志物的示例。例如,在一些情况下,分子生物标志物可以是蛋白质表达水平。例如,特异性生物标志物可以是雌激素受体(ER)、孕激素受体(PR)或人表皮生长因子受体(HER2)。这些支柱生物标志物对乳腺癌是特异性的。它们对于乳腺癌预后最重要的生物标志物,并且是靶向治疗的基础。ER和HER2最常见地分别与癌症治疗药物它莫昔芬和赫塞汀相关联。可以针对这两种生物标志物对患者进行检测,以确定对这些治疗的适用性。

[0260] 在上面描述的示例中,输入图像被分类为是否与ER生物标志物相关联。替选地,该方法可以用于确定输入图像图块是否与HER2生物标志物、PR生物标志物、EGFR生物标志物(与肺腺癌相关联)或MSI(与结肠腺癌相关联)相关联。各种分子生物标志物可以用于将某些癌症分类成诸如乳腺癌或结肠直肠癌的类。例如,乳腺癌具有五种不同的分子“亚型”,每种亚型基于ER、PR和HER2的状态确定。例如,如果ER、PR和HER2均为负,则分子亚型为“基底样型”。因此,通过确定存在或缺乏多个分子生物标志物,可以预测分子亚型。“分子亚型”是基于一个或一组生物标志物的存在或缺乏,或者一些情况下一个或一组生物标志物的水平对癌症的特定类型进行分类的方式。

[0261] 该分类方法可以用于检测其他各种其他生物标志物。例如,抗原Ki-67也被越来越多地被检测为用于指示癌症侵袭性的细胞增殖的标志物。因此,特异性生物标志物可以替选地是Ki-67。在例如评估HER2-和ER+乳腺癌的肿瘤增殖时,基于Ki67核抗原的IHC染色的标记指数可以与其他IHC标志物一起用作分级方案中有丝分裂计数的替选物。Ki-67可以针对治疗决策提供附加信息,例如是否需要辅助化疗。在各种研究中,Ki-67被示出为存活的有力预测物。例如,PREDICT是在线工具,其示出针对早期浸润性乳腺癌的不同治疗可以如何提高手术之后的生存率。在将Ki67作为预后标志物的情况下,PREDICT模型性能被改善。解释IHC染色的Ki67切片的人工评分方法包括对随机选择的感兴趣区域例如肿瘤外围中的浸润细胞进行计数,并且确定Ki67染色相对于所有浸润性肿瘤细胞的百分比。与上面描述的常规分子谱分析技术类似,该处理是劳动密集的,易于出现人为错误,并且对观察者间/观察者内开放。例如,通过根据H&E图像预测Ki67指数,可以使这一处理变得更短,并且潜在地提高准确度。

[0262] 该分类方法可以提供与癌症的诊断、治疗和/或预后相关的特定生物标志物的自动分析。特异性生物标志物可以是突变状态、受体状态或拷贝数变化等。在所描述的示例中,尽管可以使用其他图像或输入,但是分析根据全切片H&E图像执行。

[0263] 生物标志物可以是分子或者与一个或多个分子相关联的特性,例如诸如特定分子的量。在一些情况下,生物标志物是与特异性癌症治疗相关联的分子。生物标志物可以是临床上可执行的基因改变。在该示例中,特定生物标志物是ER生物标志物。其他生物标志物的示例将在下面更详细地描述。

[0264] 替选地,分类预测可以基于其他性质。例如,在一些示例中,分类方法对图像是否包含癌细胞进行分类。

[0265] 尽管在上面描述的示例中,使用了利用苏木精和伊红染色剂染色的组织切片的图像,也可以使用所描述的方法处理使用各种方法获得的各种类型的组织图像。例如,替选地,可以将已经经过免疫组织化学(IHC)染色的组织切片的图像作为输入。IHC染色包括选择性地识别组织切片的细胞中的抗原。抗体与生物组织中的抗原特异性地结合。染色允许抗体-抗原的相互作用的可视化。例如,使用显色免疫组织化学(CIH),抗体与可以催化显色

反应的酶共轭。

[0266] 在上面描述的示例中,执行二元分类法。然而,在备选示例中,分类方法也可以用于多向分类。在多向分类中,分类模型针对多个类中的每个类应用第一分类和第二分类,其中第二分类被配置成比第一分类具有针对该类的更高的真正率。给定来自验证数据集的结果,针对类的真正率被确定为该类中的该方法正确地预测为处于该类中的实例的比例。

[0267] 现在将描述执行分类成三类的示例。将三元分类器模型应用于输入。针对第一类的第一分类包括将针对第一类的输出与第一预测阈值进行比较。针对第一类的第二分类包括将针对第一类的输出与第二预测阈值进行比较。第一预测阈值高于第二预测阈值。针对第二类的第一分类包括将针对第二类的输出与第三预测阈值进行比较。针对第二类的第二分类包括将针对第二类的输出与第四预测阈值进行比较。第三预测阈值高于第四预测阈值。针对第三类的第一分类包括将针对第三类的输出与第五预测阈值进行比较。针对第三类的第二分类包括将针对第三类的输出与第六预测阈值进行比较。第五预测阈值高于第六预测阈值。第一预测阈值被设置为第一类根据验证数据集结果针对不处于第一类中(即处于第二类或第三类中)的输入的最高输出概率值。第二预测阈值被设置为第一类根据验证数据集结果针对处于第一类中的输入的最低概率值。因此,如从针对验证数据集运行分类器模的结果所确定的,第一预测阈值T1被设置为针对第一类的特异度阈值,并且第二预测阈值T2被设置为针对第一类的灵敏度阈值。第三预测阈值被设置为第二类根据验证数据集结果针对不处于第二类中的输入的最高输出概率值。第四预测阈值被设置为第二类根据验证数据集结果针对处于第二类中的输入的最低概率值。第五预测阈值被设置为第三类根据验证数据集结果针对不处于第三类中的输入的最高输出概率值。第六预测阈值被设置为第三类根据验证数据集结果针对处于第三类中的输入的最低概率值。

[0268] 现在将描述执行分类成三类的备选示例。针对第一类的第一分类包括应用第一二元分类器和使用第一预测阈值,并且针对第一类的第二分类包括应用第一分类器和使用第二预测阈值,其中,第一预测阈值高于第二预测阈值。对第一二元分类器进行训练,以确定输入是处于第一类中(正类)还是不处于第一类中(负类)。针对第二类的第一分类包括应用第二二元分类器和使用第三预测阈值,并且针对第二类的第二分类包括应用第二分类器和使用第四预测阈值,其中,第三预测阈值高于第四预测阈值。第二二元分类器确定输入是处于第二类中(正类)还是不处于第二类中(负类)。针对第三类的第一分类包括应用第三二元分类器和使用第五预测阈值,并且针对第三类的第二分类包括应用第三分类器和使用第六预测阈值,其中,第五预测阈值高于第六预测阈值。第二二元分类器确定输入是处于第三类中(正类)还是不处于第三类中(负类)。每个分类器针对训练数据集单独地进行训练,并且如结合第一示例所描述的对阈值进行选择。

[0269] 现在将描述执行分类成三类的备选示例。针对第一类的第一分类包括应用第一二元分类器和使用第一预测阈值,并且针对第一类的第二分类包括应用第二二元分类器和使用第二预测阈值。对第一二元分类器和第二二元分类器进行训练,以确定输入是处于第一类中(正类)还是不处于第一类中(负类)。针对第二类的第一分类包括应用第三二元分类器和使用第三预测阈值,并且针对第二类的第二分类包括应用第四二元分类器和使用第四预测阈值。对第三二元分类器和第四二元分类器进行训练,以确定输入是处于第二类中(正类)还是不处于第二类中(负类)。针对第三类的第一分类包括应用第五二元分类器和使用

第五预测阈值,并且针对第三类的第二分类包括应用第六二元分类器和使用第六预测阈值。对第五二元分类器和第六二元分类器进行训练,以确定输入是处于第三类中(正类)还是不处于第三类中(负类)。针对训练数据集单独地训练与每个类对应的三个或更多个分类器,并且如上面结合第二示例描述的针对每个类选择用于第一分类和第二分类的分类器。对于每个类,针对两个分类器的正类是相同的。

[0270] 图14是分类方法的示意图。该方法包括部署不仅输出二元预测值而且基于其针对测试数据集的先前表现输出是否确定预测正确的指示符的预测医疗设备(PMD)。在PMD确定的情况下,针对实例给出预测。在PMD不确定的情况下,不给出预测,并且实例将继续按照标准护理进行。

[0271] 在S1301中,获得要分析的输入医疗样本或数据。在S1302中,将输入输入至PMD。在S1303中,输出预测值(1或0)以及确定性。如果第一分类和第二分类一致,则确定性为1,并且方法移动至S1305。绕过护理标准分析,提供分析(预测值)。如果第一分类和第二分类不一致,则确定性为0,并且方法移动至S1304。然后执行护理标准分析,其中可能包括手动检查。该分析可能是耗时的或昂贵的。一旦在S1305或S1304中提供了分析,方法就会移动至S1306并且执行后续处理。

[0272] 图15示出了根据实施方式的分类系统1的示意图。系统1包括输入11、处理器3、工作存储器9、输出13和存储装置7。系统1接收输入图像数据并且生成输出。输出可以包括诊断信息。特别地,输出可以是输入图像是否与生物标志物相关联的指示。

[0273] 系统1是计算系统,例如诸如通用计算机的终端用户系统、或服务器。

[0274] 处理器3耦接至存储装置7,并且访问包括RAM 9的工作存储器。处理器3可以包括对工作存储器中存储的代码中的指令进行响应和处理的逻辑电路系统。

[0275] 计算机程序5存储在非易失性存储器7中。由处理器3访问非易失性存储器7,并且由处理器3检索和执行所存储的代码5。特别地,体现本文描述的方法的计算机程序代码5在执行时被表示为存储在工作存储器9中的软件产品。由处理器3执行代码5将使如本文所描述的实施方式实现。

[0276] 处理器3还访问输入模块11和输出模块13。输入模块和输出模块或者输入接口和输出接口11、13可以是单个部件,或者可以划分成单独的输入接口11和单独的输出接口13。

[0277] 输入模块11连接至输入15以用于接收数据。输入15可以是用于从外部存储介质或通过通信网络接收数据的接收器。替选地,输入15可以包括硬件,例如图像捕获设备。替选地,输入15可以从存储的图像文件读取数据,该图像文件可以存储在系统上,或者可以存储在单独的存储介质例如软盘、CD ROM、磁装置或可编程存储装置上。

[0278] 输出17连接至输出模块13。输出17可以包括硬件,例如视觉显示器。替选地,输出可以是用于将数据发射到外部存储介质或通过通信网络发射数据的发射器。另外,输出17可以将数据写入存储的图像文件中,该图像文件可以存储在系统上,或者可以存储在单独的存储介质例如软盘、CD ROM、磁装置或可编程存储装置上。

[0279] 存储装置7通信上耦接至处理器3。存储装置7可以包含在由处理器3执行代码5时由代码5使用的数据。如所示出的,存储装置7是包含在装置中的本地存储器。然而,替选地,存储装置7可以全部或部分位于远程,例如,使用可以经由通信网络(例如互联网)远程访问的基于云的存储器。代码5也存储在存储装置7中。代码5在执行时被放置在工作存储器9中。

[0280] 系统1可以与用于输入和输出数据的硬件位于共同的系统中。替选地,系统1可以是远程系统1,该远程系统1接收从单独的单元(例如,图像捕获装置)发送的图像数据,并且将输出数据发送至另一单独的单元(例如,包括屏幕的用户计算机)。例如,该系统可以在云计算系统上实现,该云计算系统接收并发送数据。

[0281] 尽管在所描述的系统中,使用了位于装置中的单个处理器3,但是系统可以包括两个或更多个处理器,所述处理器可以位于同一系统中或者位于远程,被配置成执行处理的不同部分并且在它们之间发送数据。例如,系统可以包括图形处理单元(GPU)和通用中央处理单元(CPU)。各种操作由GPU实现,而其他操作由CPU实现。例如,矩阵运算可以由GPU执行。

[0282] 适用用于将软件加载至存储器中和将数据存储在存储单元7中的常规流程。代码5可以嵌入原始设备中,或者可以在生产后作为整体或部分被提供。例如,代码可以作为计算机程序产品整体被引入,其可以是以下载的形式,或者可以经由计算机程序存储介质例如光盘被引入。替选地,可以通过更新或插件对现有的对话管理器软件做出修改,以提供所描述的実施方式的特征。

[0283] 训练分类模型的方法和使用经训练的分类模型的分类方法可以在同一系统上执行。例如,训练分类模型的方法和使用经训练的分类模型的分类方法可以在如结合图15所描述的系统上执行,其中经训练的分类模型一旦被训练就存储在存储器7中。替选地,可以在第一系统上执行训练分类模型的方法,并且将经训练的分类模型存储在第二系统中,使得在第二系统上执行分类方法。

[0284] 虽然将理解,所描述的實施方式适用于任何计算系统,但是图15中示出的示例计算系统提供了能够使如本文所描述的實施方式生效的手段。在使用时,系统1通过数据输入11接收数据。在处理器3上执行的程序5以参照上面的图已经描述的方式通过输出13输出数据。

[0285] 尽管已经描述了特定的實施方式,但是这些實施方式仅通过示例的方式呈现,并且不旨在限制本发明的范围。事实上,本文中描述的新方法和设备可以以多种其他形式体现;此外,可以对本文中描述的方法和设备的形式做出各种省略、替换和变化。

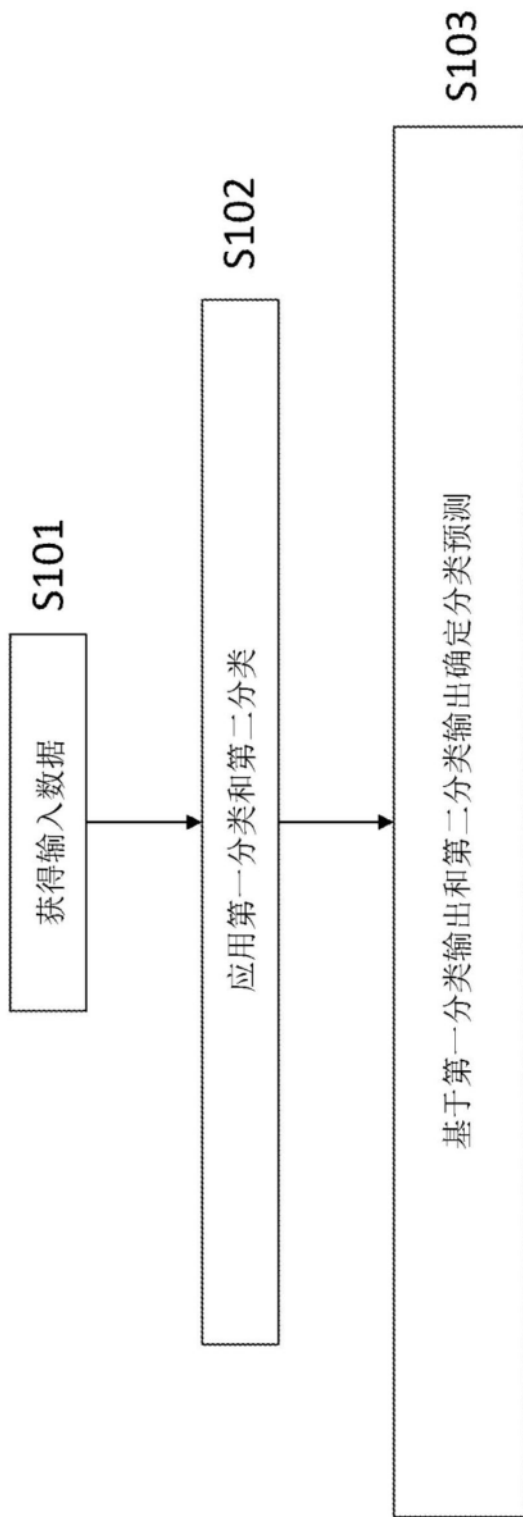


图1

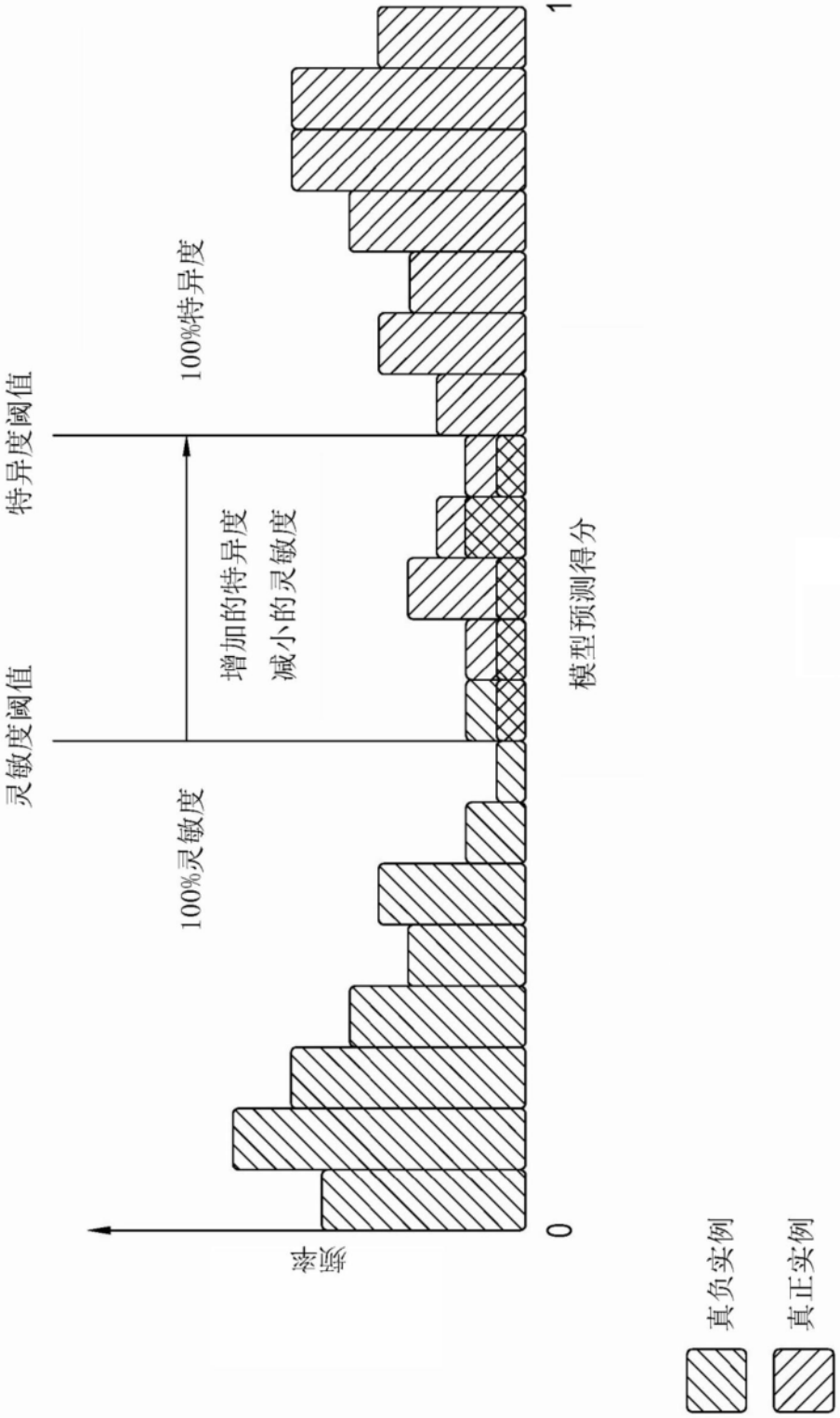


图2(a)

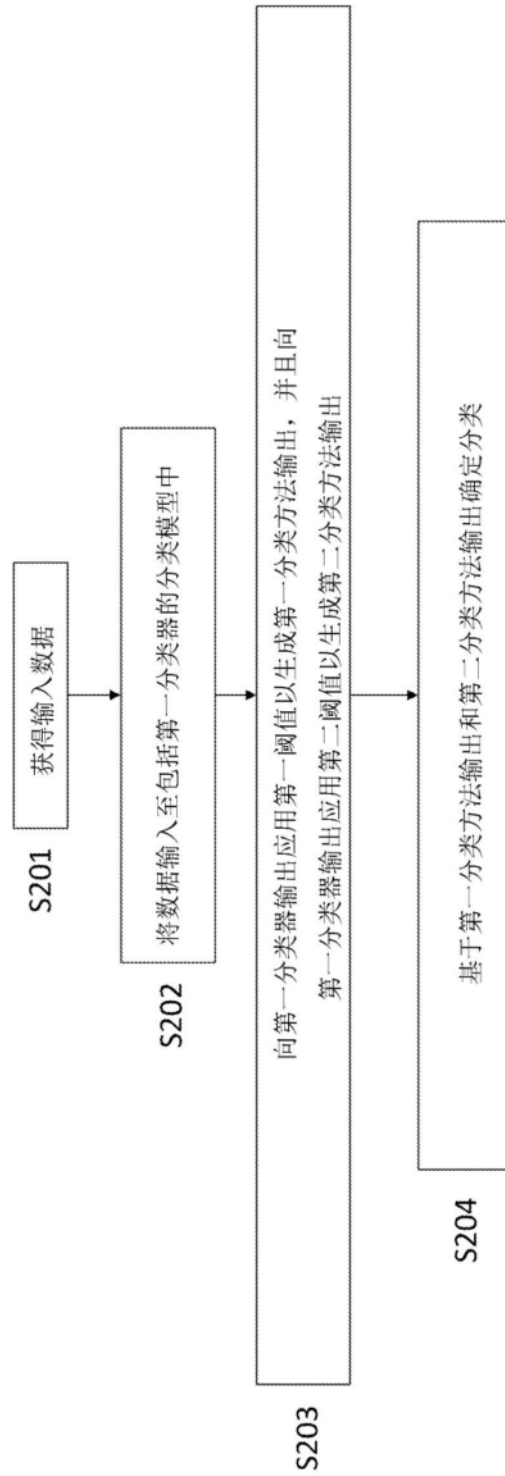


图2(b)

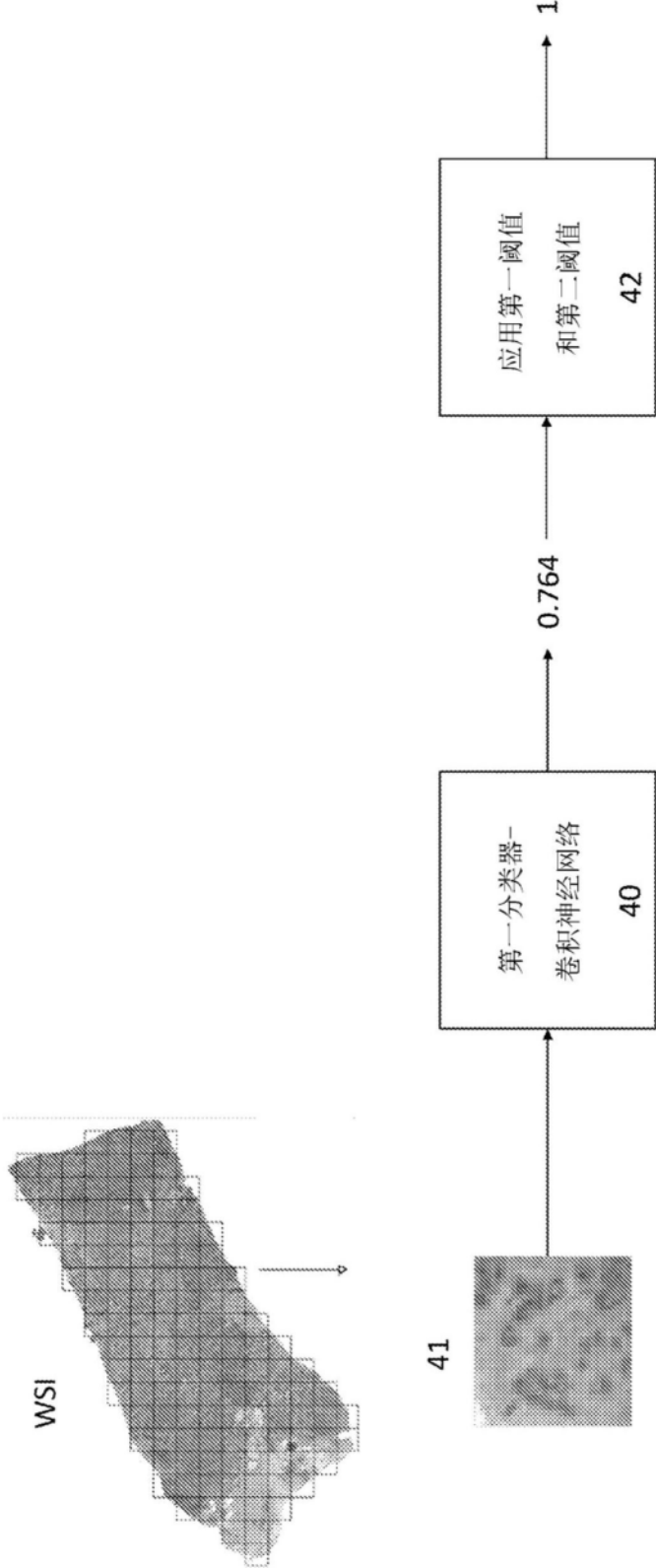


图3

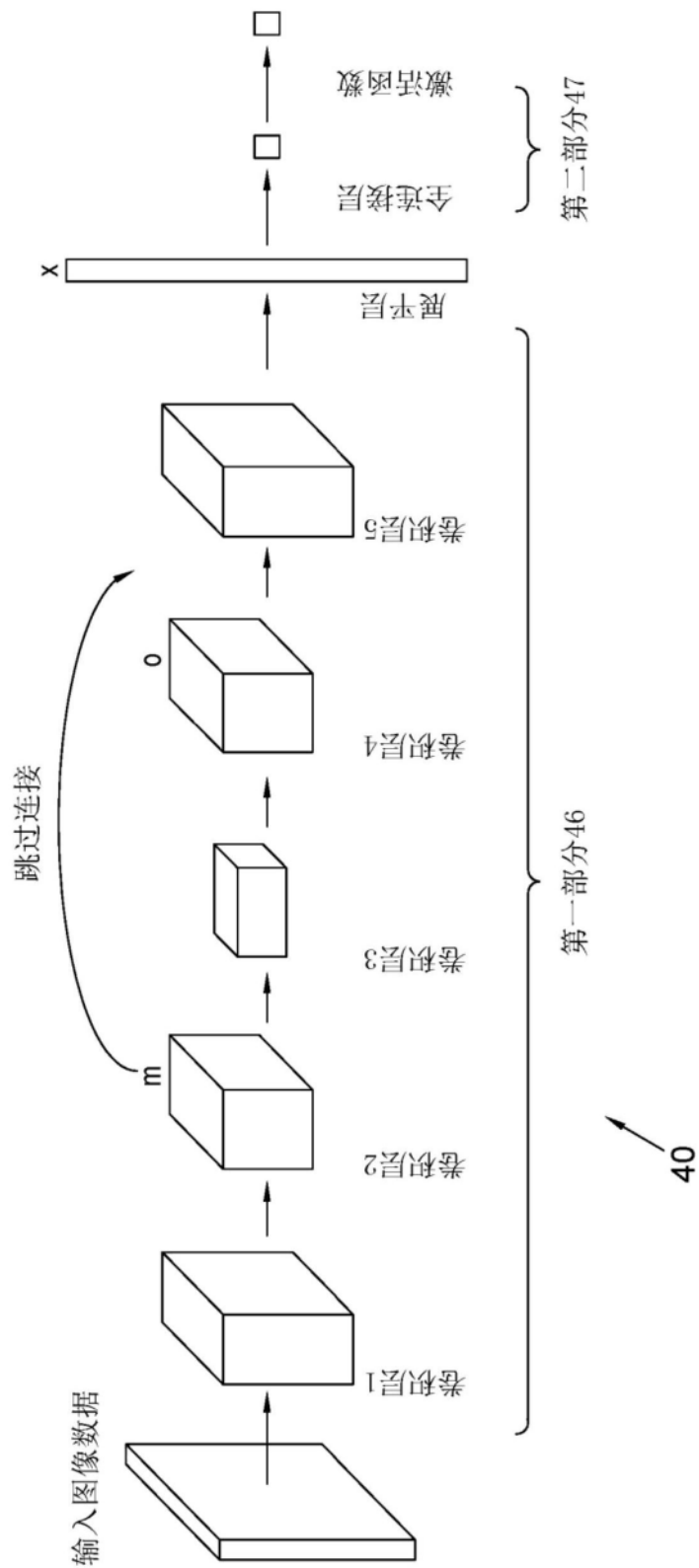


图4

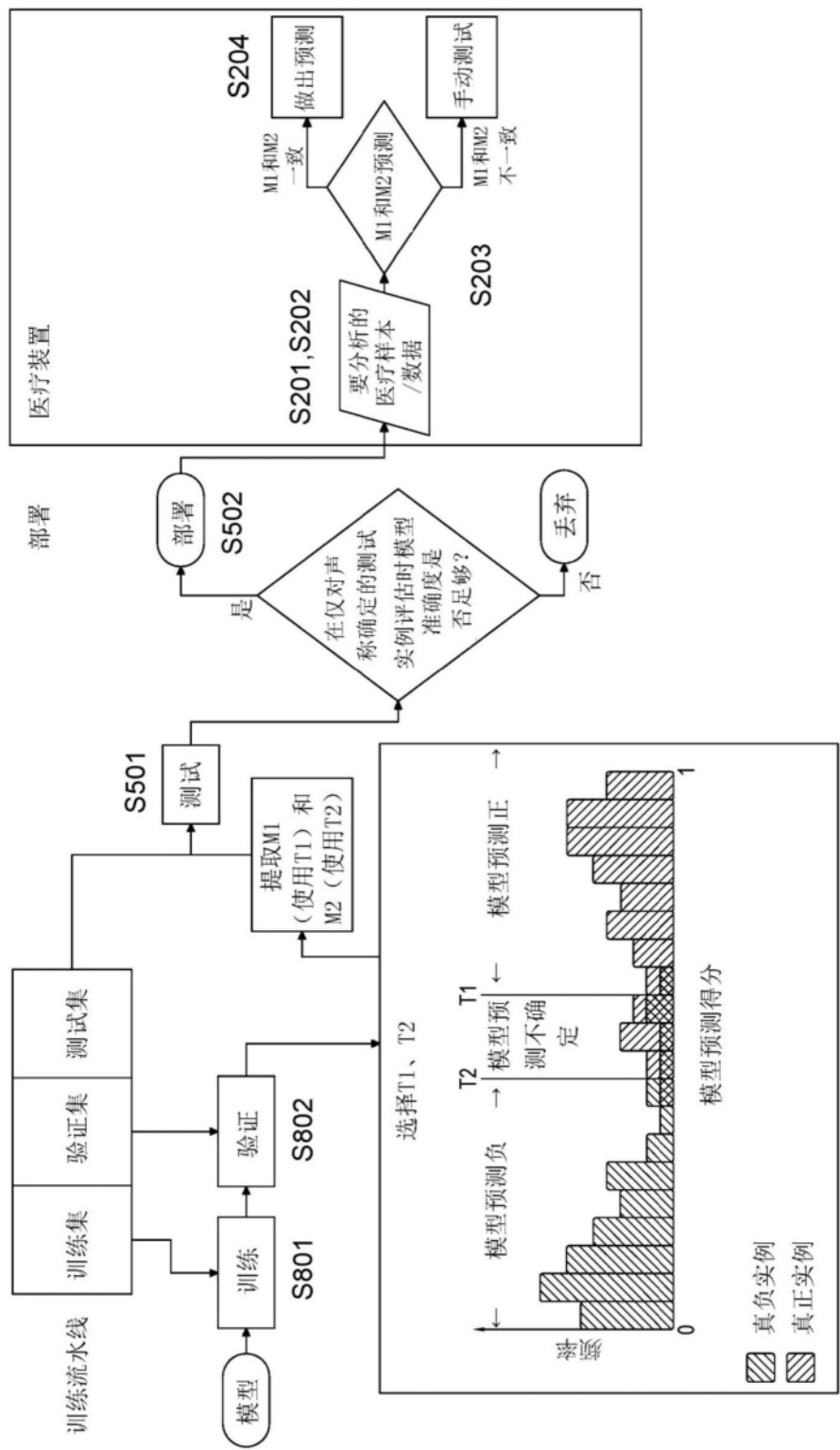


图5

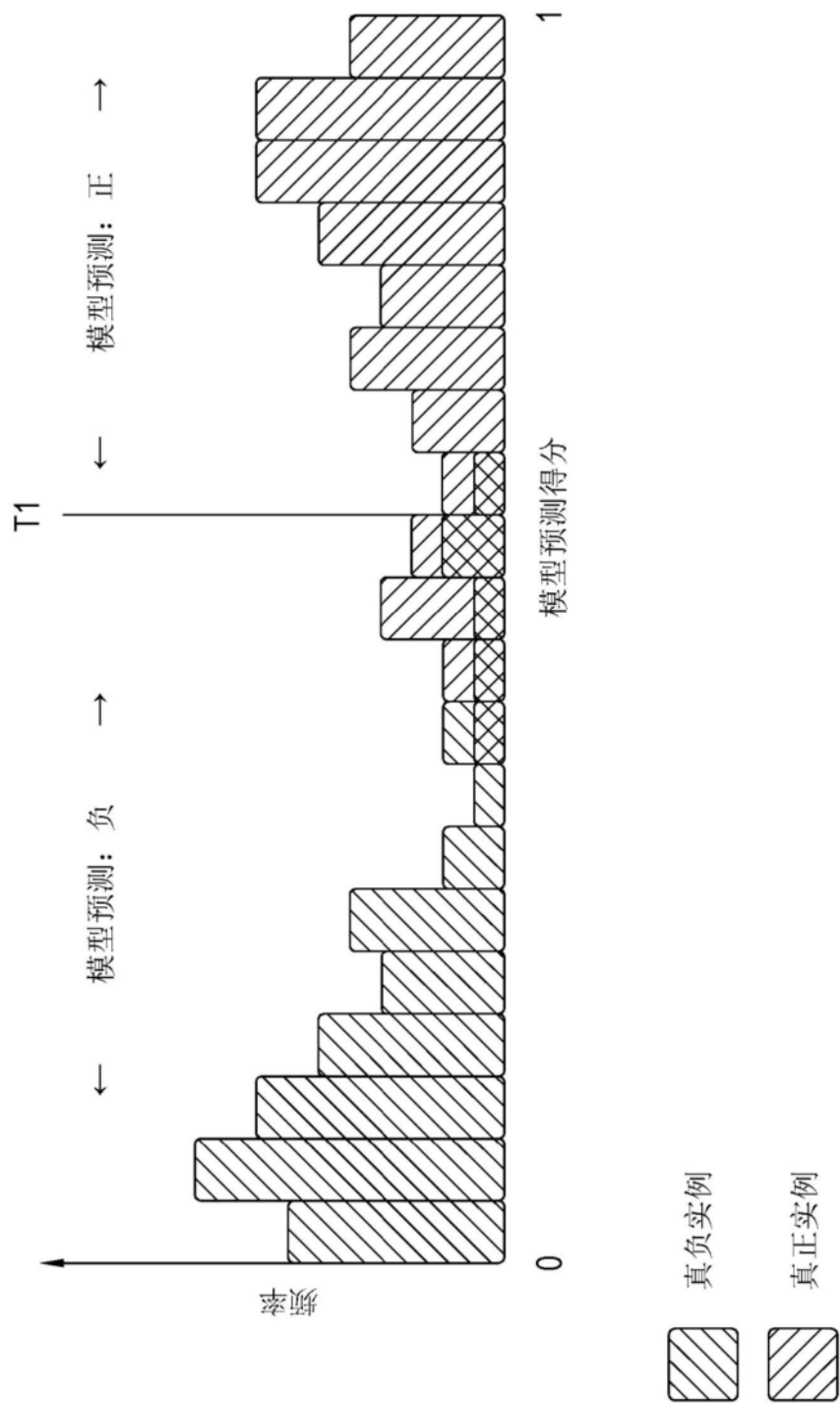


图6



图7

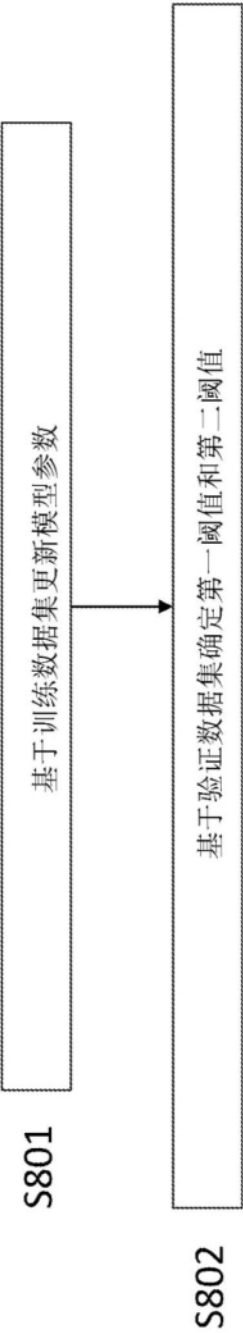


图8(a)

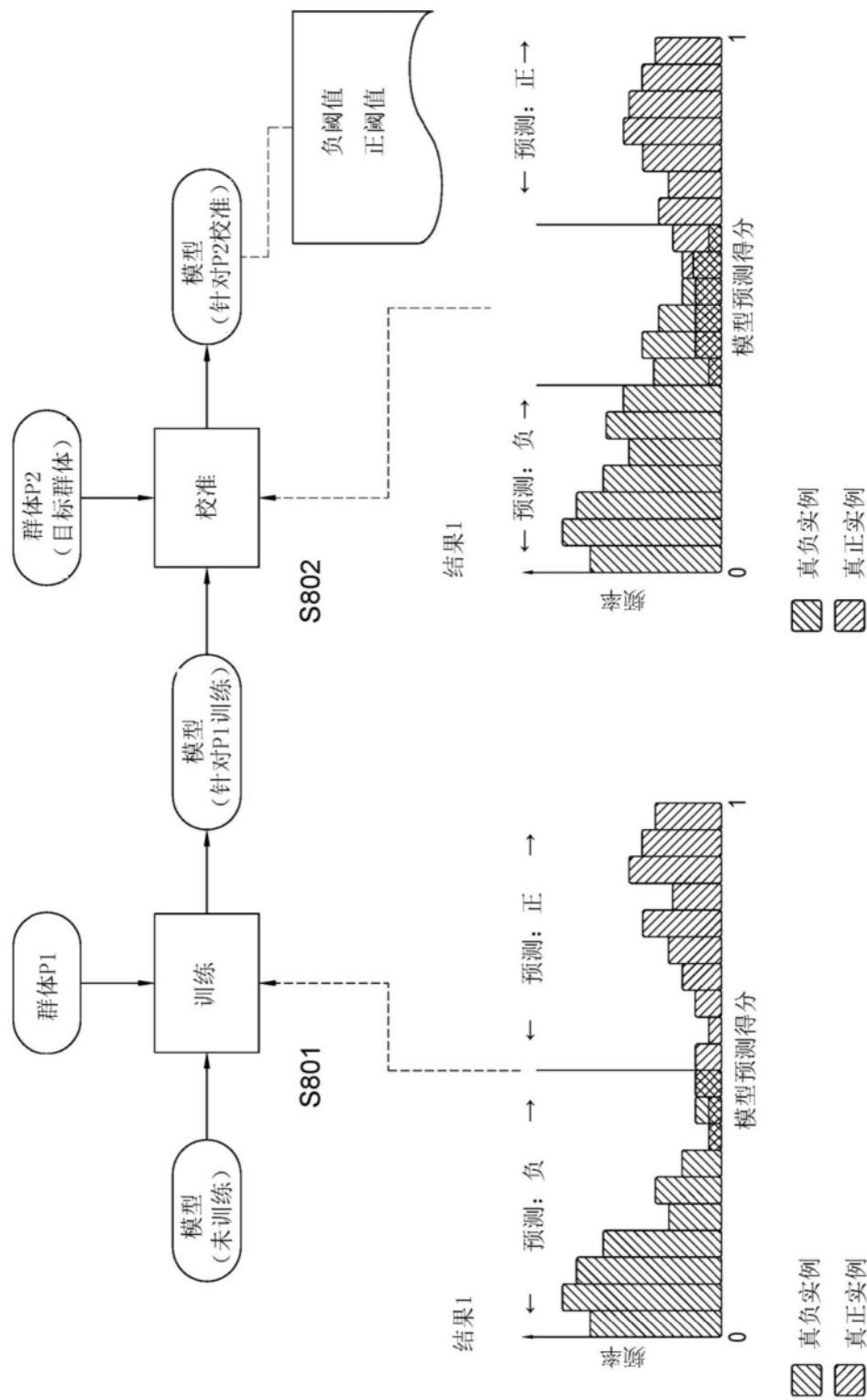


图8(b)

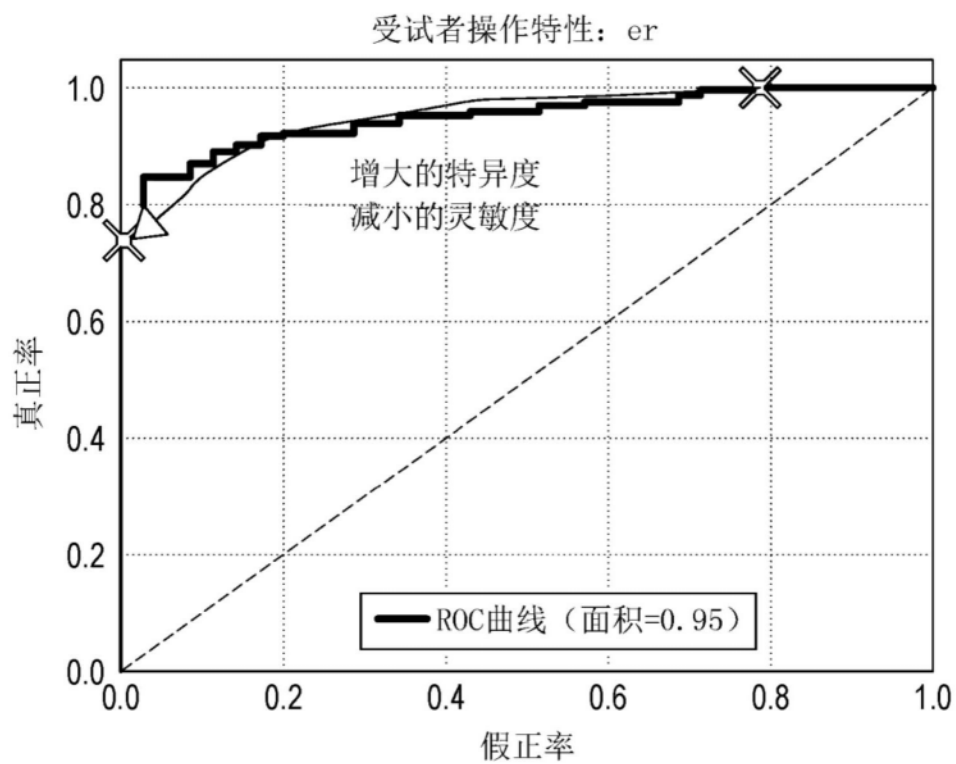


图9

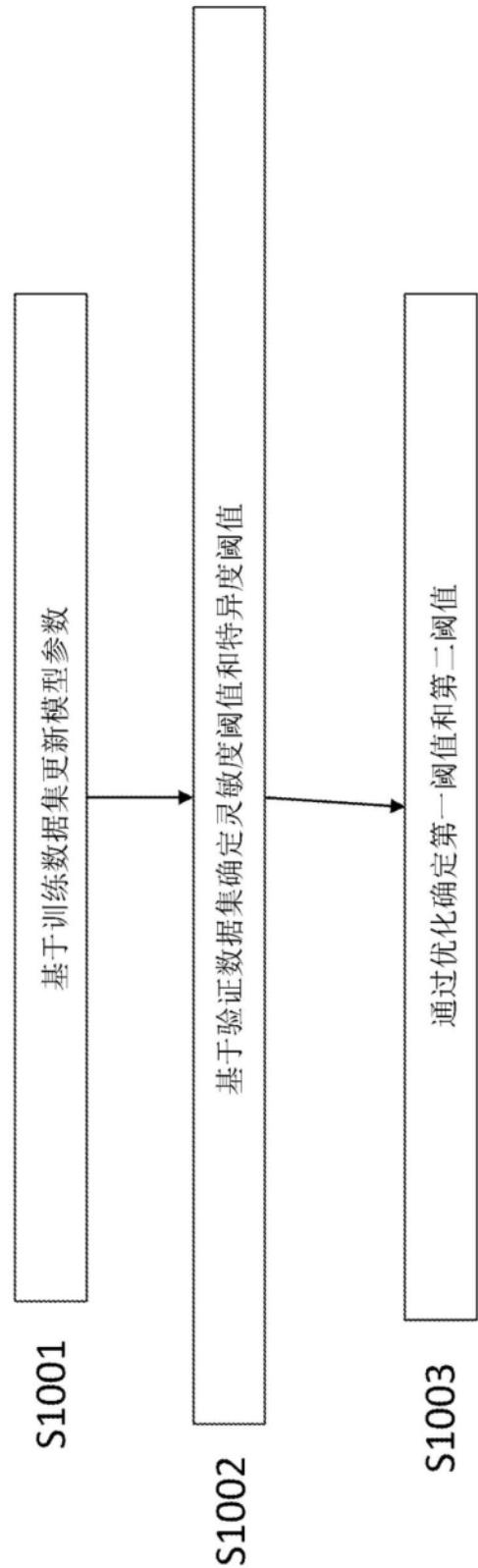


图10(a)

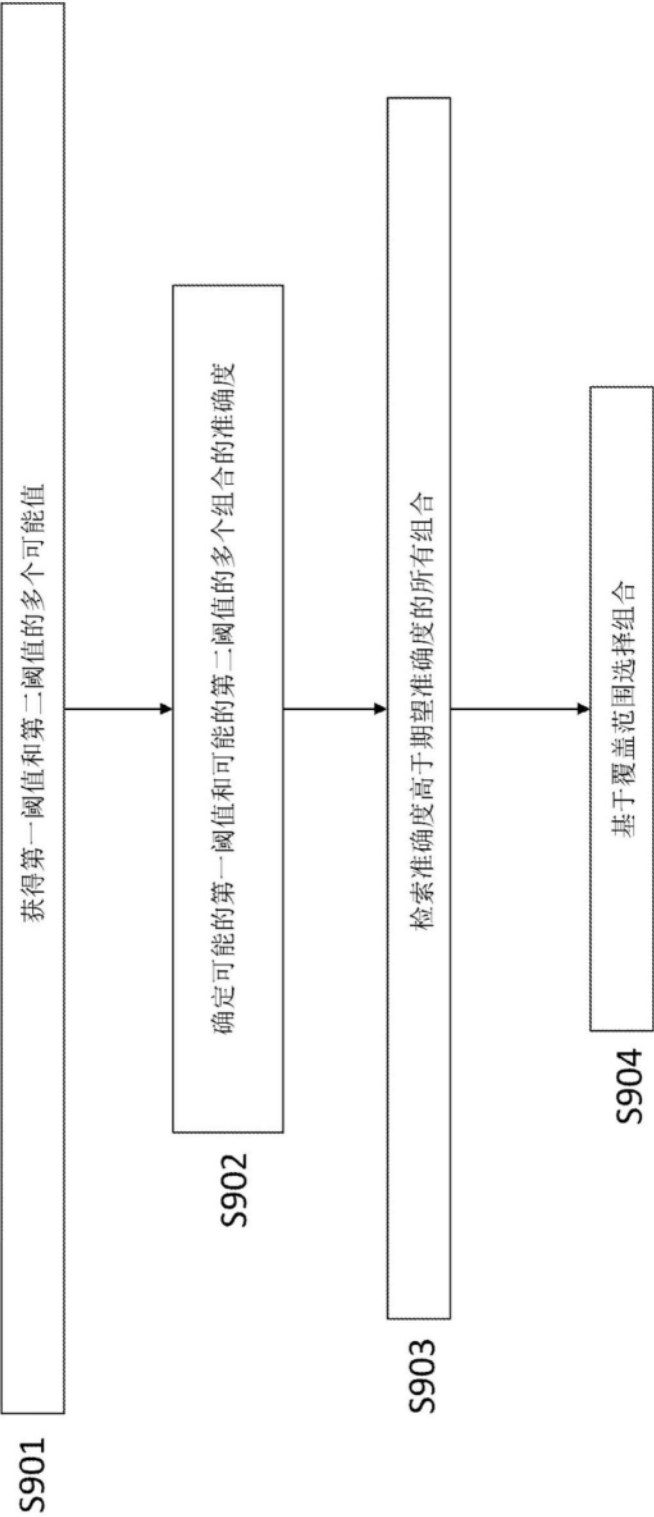


图10(b)

T1 T2	0.425	0.455	0.475	0.500	0.525	0.550	0.575	0.600	0.625	0.650
0.400	0.34	0.34	0.38	0.61	0.72	0.85	0.91	0.96	0.98	0.99
0.425		0.31	0.37	0.58	0.74	0.83	0.89	0.94	0.96	0.97
0.450			0.33	0.53	0.72	0.81	0.87	0.87	0.95	0.95
0.475				0.47	0.69	0.75	0.80	0.83	0.92	0.93
0.500					0.60	0.65	0.75	0.82	0.82	0.84
0.525						0.42	0.61	0.63	0.63	0.66
0.550							0.35	0.37	0.39	0.41
0.575								0.37	0.40	0.40
0.600									0.37	0.38
0.625										0.32

图10(c)

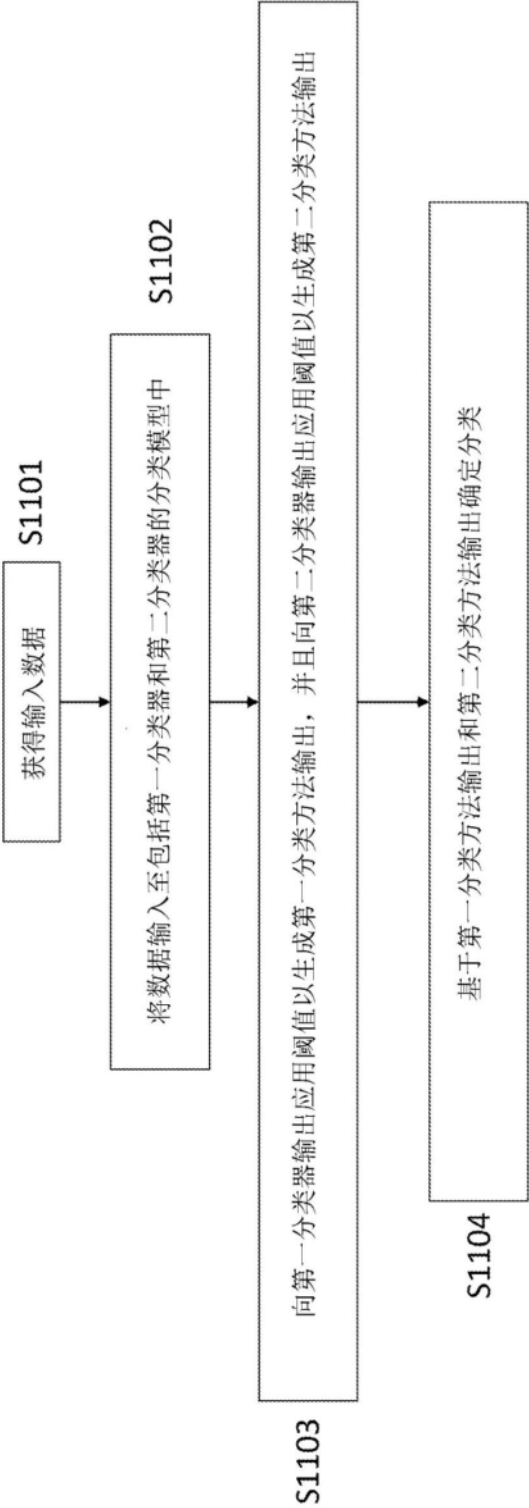


图11(a)

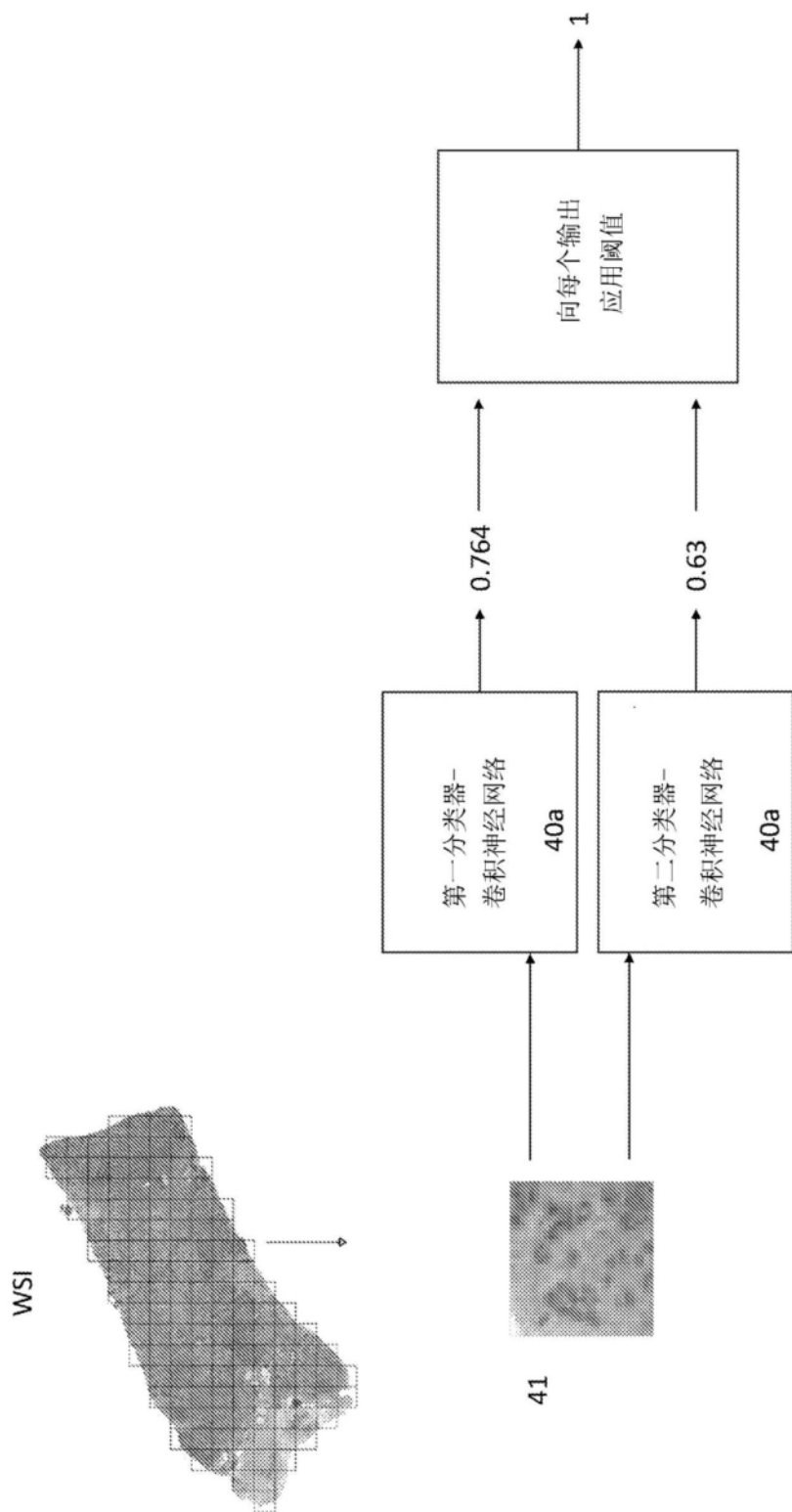


图11(b)

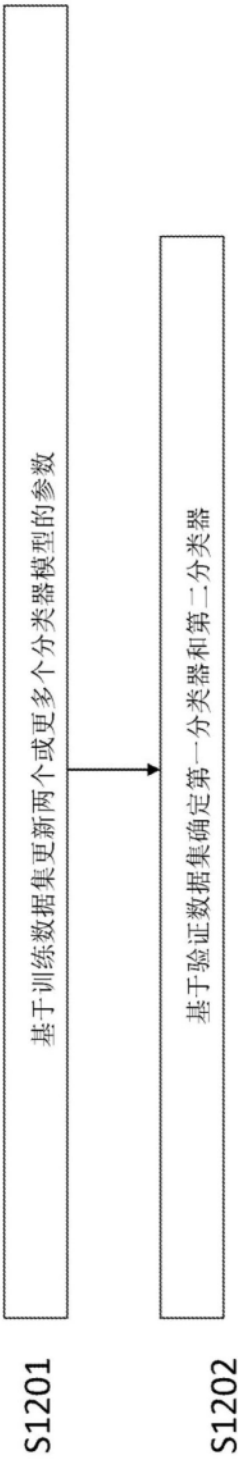


图12(a)

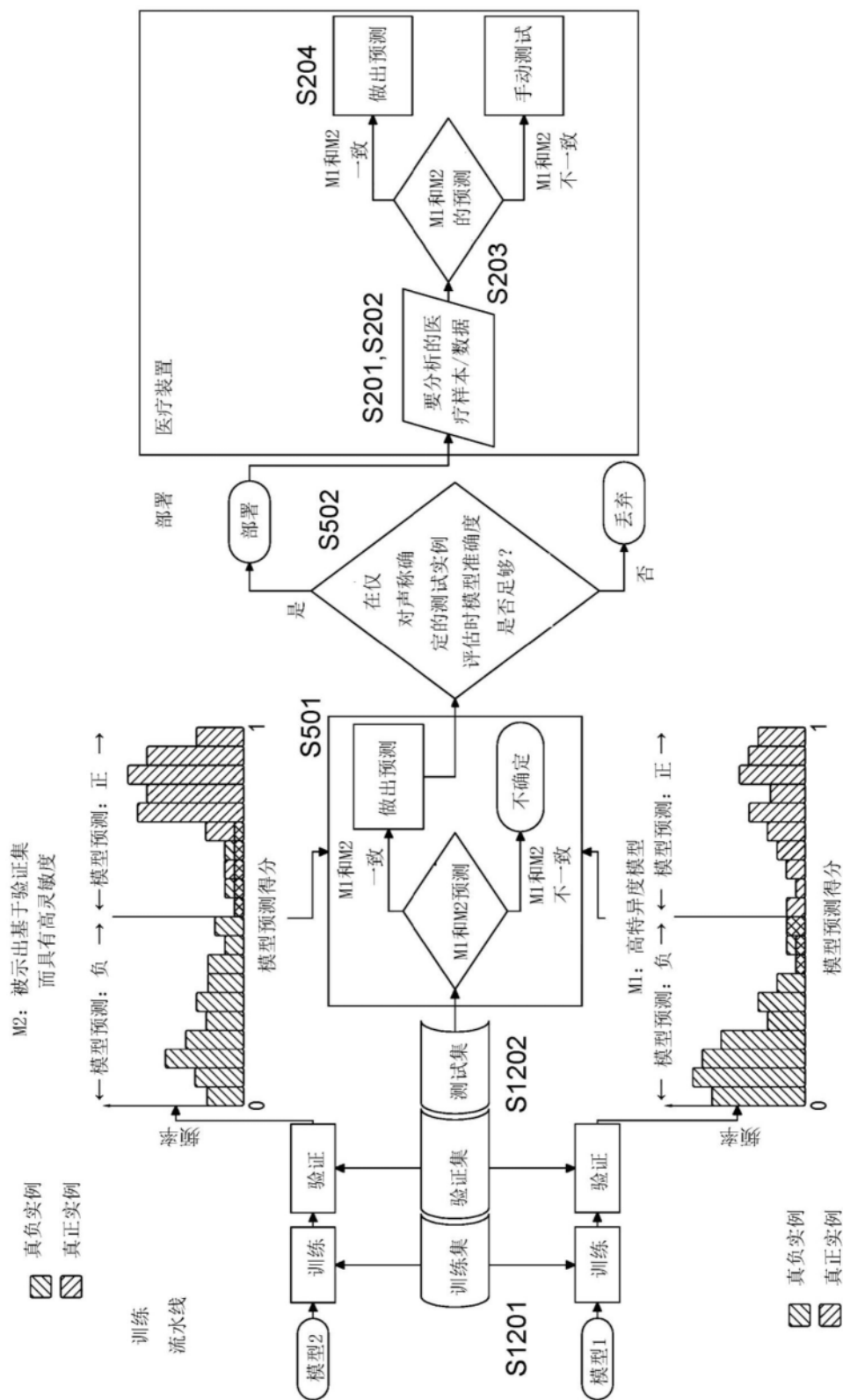


图12(b)

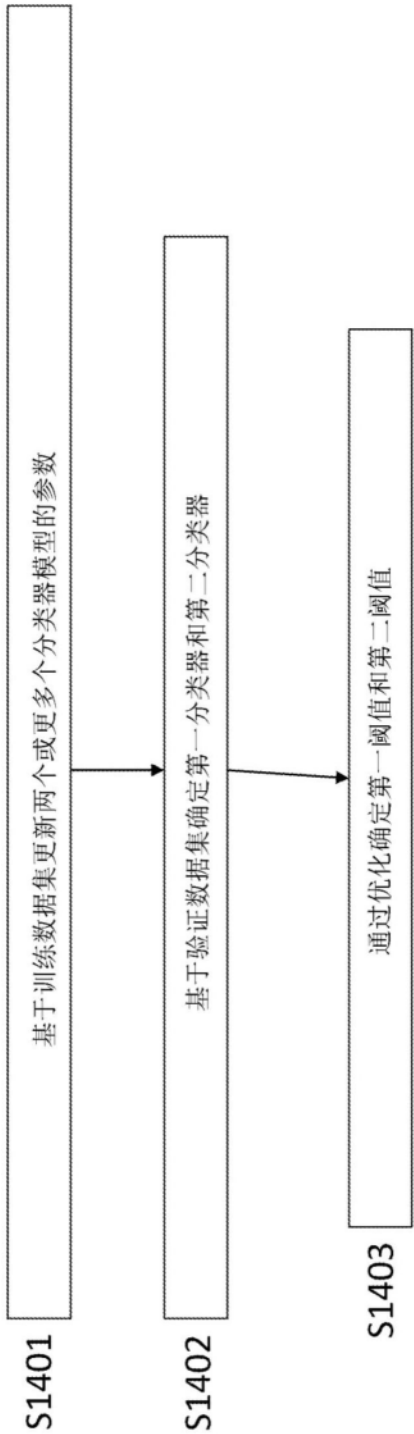
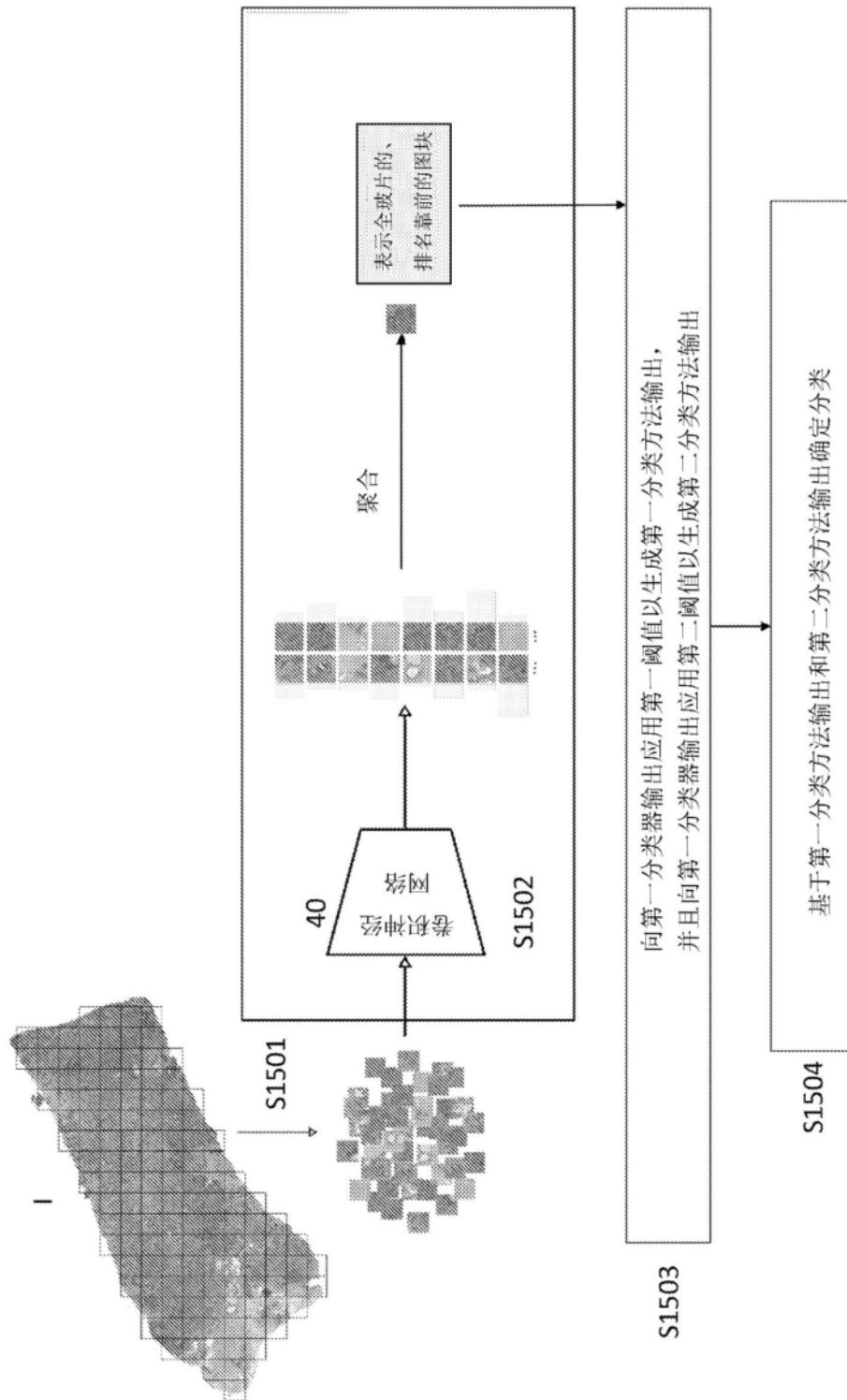


图12(c)



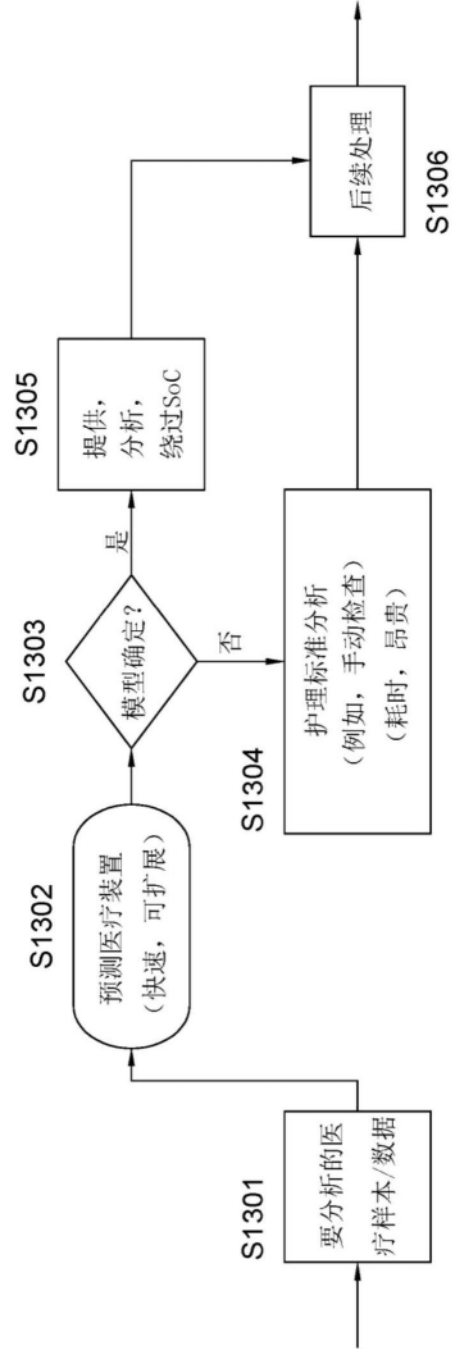


图14

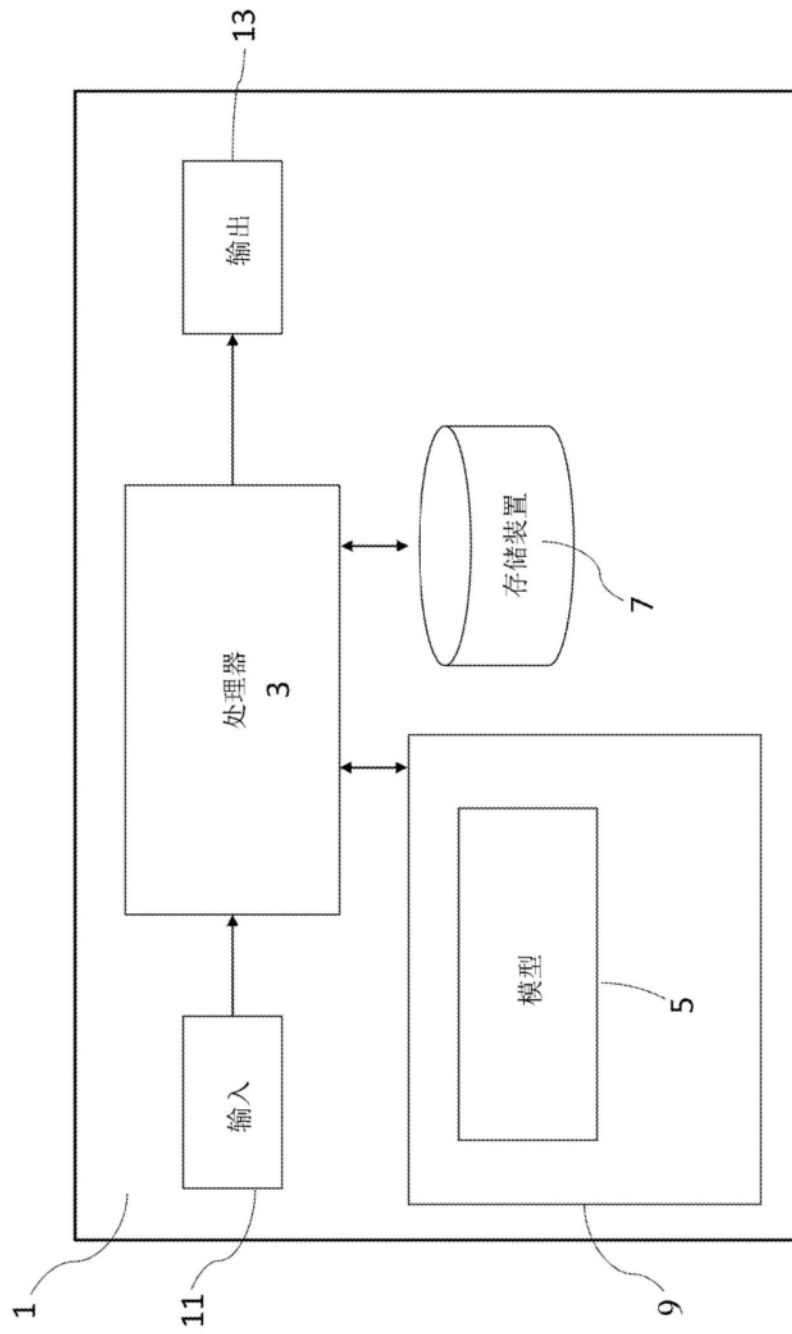


图15

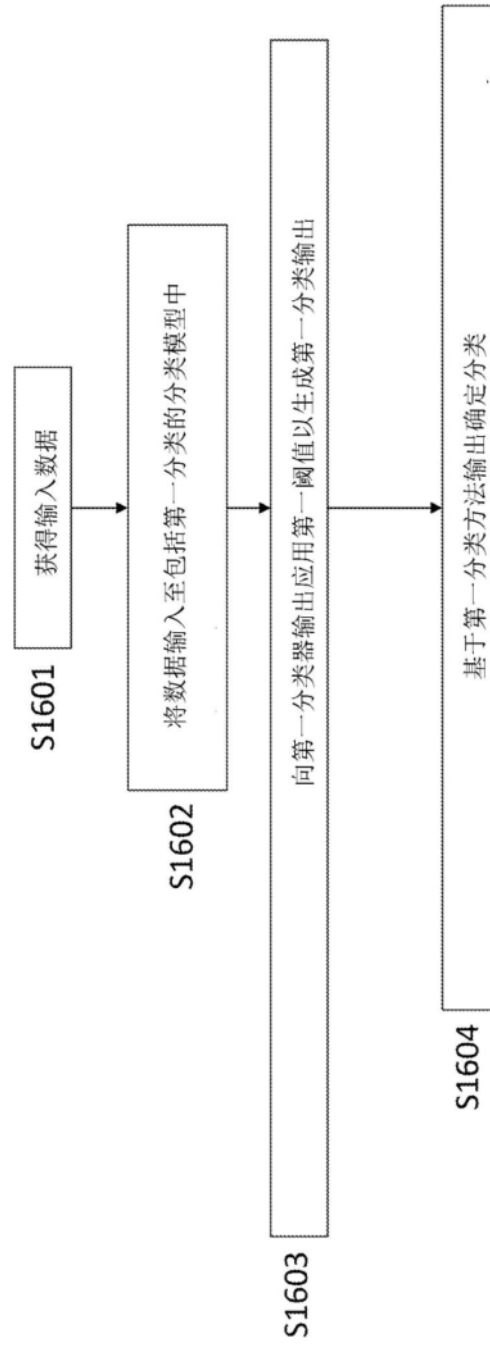


图16

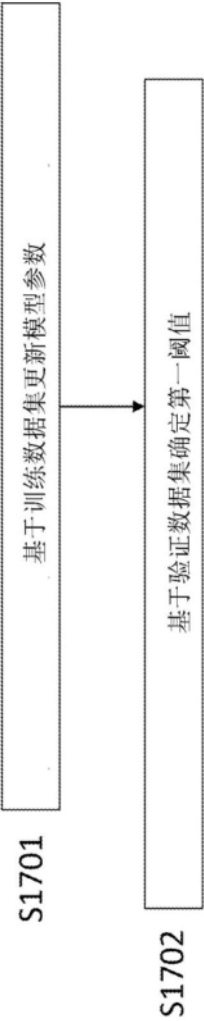


图17