



(12) 发明专利申请

(10) 申请公布号 CN 114445651 A

(43) 申请公布日 2022. 05. 06

(21) 申请号 202111581998.5

(22) 申请日 2021.12.22

(71) 申请人 天翼云科技有限公司

地址 100007 北京市东城区青龙胡同甲1号、3号2幢2层205-32室

(72) 发明人 朱锦程

(51) Int. Cl.

G06V 10/764 (2022.01)

G06V 20/70 (2022.01)

G06V 10/82 (2022.01)

G06V 10/774 (2022.01)

G06V 10/26 (2022.01)

G06K 9/62 (2022.01)

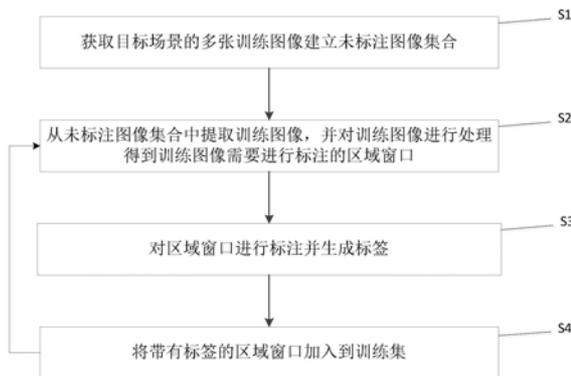
权利要求书2页 说明书7页 附图2页

(54) 发明名称

一种语义分割模型的训练集构建方法、装置及电子设备

(57) 摘要

本发明公开了一种语义分割模型的训练集构建方法,该方法包括:获取目标场景的多张训练图像建立未标注图像集合;从未标注图像集合中提取训练图像,并对训练图像进行处理得到训练图像需要进行标注的区域窗口;对区域窗口进行标注并生成标签;将带有标签的区域窗口加入到训练集,并返回从未标注图像集合中提取训练图像的步骤,直至训练集满足预设训练集要求。本发明通过自动筛选不确定性较大的区域并生成区域窗口进行标注,代替传统的整张图像的标注,极大地减少了人工标注的人力成本和时间,加快了网络对于特定场景分割任务的训练集构建速度,提高了训练集的构建效率。



1. 一种语义分割模型的训练集构建方法,其特征在于,包括:
 - 获取目标场景的多张训练图像建立未标注图像集合;
 - 从所述未标注图像集合中提取训练图像,并对所述训练图像进行处理得到训练图像需要进行标注的区域窗口;
 - 对所述区域窗口进行标注并生成标签;
 - 将带有标签的区域窗口加入到训练集,并返回从所述未标注图像集合中提取训练图像的步骤,直至所述训练集满足预设训练集要求。
2. 根据权利要求1所述的语义分割模型的训练集构建方法,其特征在于,所述对所述训练图像进行处理得到每张训练图像需要进行标注的区域窗口,包括:
 - 对所述训练图像进行特征处理得到特征区域图像;
 - 基于所述特征区域图像确定需要进行标注的区域窗口。
3. 根据权利要求2所述的语义分割模型的训练集构建方法,其特征在于,所述对所述训练图像进行特征处理得到特征区域图像,包括:
 - 通过预设算法提取所述训练图像上的多个特征点图像;
 - 将所述特征点图像分别代入多个预设模型,得到特征点图像的像素分类结果;
 - 对所述像素分类结果的一致性进行判断;
 - 选取一致性判断结果未达到预设目标的特征点图像作为特征区域图像。
4. 根据权利要求2所述的语义分割模型的训练集构建方法,其特征在于,所述基于所述特征区域图像确定需要进行标注的区域窗口,包括:
 - 筛选所述特征区域图像中信息熵最大的像素点坐标;
 - 将所述像素点坐标作为中心点对所述训练图像进行划分,得到需要进行标注的区域窗口。
5. 根据权利要求4所述的语义分割模型的训练集构建方法,其特征在于,所述将所述像素点坐标作为中心点对所述训练图像进行划分,得到需要进行标注的区域窗口,包括:
 - 获取所述训练图像的边界信息;
 - 根据所述像素点坐标和预设的窗口尺寸确定预选区域;
 - 根据所述边界信息对所述预选区域进行调整,得到区域窗口。
6. 根据权利要求1所述的语义分割模型的训练集构建方法,其特征在于,所述对所述区域窗口进行标注并生成标签,包括:
 - 将需要进行标注的区域窗口加入到待处理集合;
 - 对所述待处理集合内的区域窗口进行人工标注,并生成标签。
7. 一种语义分割模型的训练集构建装置,其特征在于,包括:
 - 获取模块,用于获取目标场景的多张训练图像构建未标注图像集合;
 - 处理模块,用于从所述未标注图像集合中提取训练图像,并对所述训练图像进行处理得到训练图像需要进行标注的区域窗口;
 - 标注模块,用于对所述区域窗口进行标注并生成标签;
 - 集合模块,用于将带有标签的区域窗口加入到训练集,并返回从所述未标注图像集合中提取训练图像的步骤,直至所述训练集达到预设的精度。
8. 根据权利要求7所述的语义分割模型的训练集构建装置,其特征在于,所述处理模

块,包括:

特征模块,用于对所述训练图像进行特征处理得到特征区域图像;

确定模块,用于基于所述特征区域图像确定需要进行标注的区域窗口。

9. 一种电子设备,其特征在于,包括:

存储器和处理器,所述存储器和所述处理器之间互相通信连接,所述存储器中存储有计算机指令,所述处理器通过执行所述计算机指令,从而执行权利要求1-6中任一项所述的语义分割模型的训练集构建方法。

10. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储有计算机指令,所述计算机指令用于使所述计算机执行权利要求1-6中任一项所述的语义分割模型的训练集构建方法。

一种语义分割模型的训练集构建方法、装置及电子设备

技术领域

[0001] 本发明涉及计算机视觉技术领域，具体涉及一种语义分割模型的训练集构建方法、装置及电子设备。

背景技术

[0002] 语义分割是现在人工智能领域的一大热门技术，该技术通过特定的神经网络可以将图像或者视频中展现的场景类别一次性的分割出来。分割出的结果不仅包含图像中存在的类别信息，同时包含了各个类别的位置、形状等信息。现如今很多应用场景中都会需要这样的信息，例如自动驾驶、医疗影像、机器人场景理解等。可以看出语义分割的应用前景十分广泛。

[0003] 高质量的语义分割数据集对于该领域的发展是至关重要的。现有的语义分割数据都是通过人工标注而来的，对图像上的每一个像素点进行类别标记最终得到分割所需的标签，这样的制作过程会消耗大量的人力和时间。尤其在庞大的数据集需求下，人工制作的成本还是非常大的。现有的主动式学习方法基于整张图像来进行筛选的主动式策略忽略了大部分数据中类别信息不对等的情况，导致一些容易被网络学习的类别部分过度标注，消耗人力成本。

发明内容

[0004] 有鉴于此，本发明实施例提供了一种语义分割模型的训练集构建方法，以解决现有的学习方法基于整张图像来进行筛选的方式导致成本高、效率低的问题。

[0005] 为达到上述目的，本发明提供如下技术方案：

[0006] 本发明实施例提供了一种语义分割模型的训练集构建方法，包括：

[0007] 获取目标场景的多张训练图像建立未标注图像集合；

[0008] 从所述未标注图像集合中提取训练图像，并对所述训练图像进行处理得到训练图像需要进行标注的区域窗口；

[0009] 对所述区域窗口进行标注并生成标签；

[0010] 将带有标签的区域窗口加入到训练集，并返回从所述未标注图像集合中提取训练图像的步骤，直至所述训练集满足预设训练集要求。

[0011] 可选的，所述对所述训练图像进行处理得到每张训练图像需要进行标注的区域窗口，包括：

[0012] 对所述训练图像进行特征处理得到特征区域图像；

[0013] 基于所述特征区域图像确定需要进行标注的区域窗口。

[0014] 可选的，所述对所述训练图像进行特征处理得到特征区域图像，包括：

[0015] 通过预设算法提取所述训练图像上的多个特征点图像；

[0016] 将所述特征点图像分别代入多个预设模型，得到特征点图像的像素分类结果；

[0017] 对所述像素分类结果的一致性进行判断；

- [0018] 选取一致性判断结果未达到预设目标的特征点图像作为特征区域图像。
- [0019] 可选的,所述基于所述特征区域图像确定需要进行标注的区域窗口,包括:
- [0020] 筛选所述特征区域图像中信息熵最大的像素点坐标;
- [0021] 将所述像素点坐标作为中心点对所述训练图像进行划分,得到需要进行标注的区域窗口。
- [0022] 可选的,所述将所述像素点坐标作为中心点对所述训练图像进行划分,得到需要进行标注的区域窗口,包括:
- [0023] 获取所述训练图像的边界信息;
- [0024] 根据所述像素点坐标和预设的窗口尺寸确定预选区域;
- [0025] 根据所述边界信息对所述预选区域进行调整,得到区域窗口。
- [0026] 可选的,所述对所述区域窗口进行标注并生成标签,包括:
- [0027] 将需要进行标注的区域窗口加入到待处理集合;
- [0028] 对所述待处理集合内的区域窗口进行人工标注,并生成标签。
- [0029] 本发明实施例还提供了一种语义分割模型的训练集构建装置,包括:
- [0030] 获取模块,用于获取目标场景的多张训练图像构建未标注图像集合;
- [0031] 处理模块,用于从所述未标注图像集合中提取训练图像,并对所述训练图像进行处理得到训练图像需要进行标注的区域窗口;
- [0032] 标注模块,用于对所述区域窗口进行标注并生成标签;
- [0033] 集合模块,用于将带有标签的区域窗口加入到训练集,并返回从所述未标注图像集合中提取训练图像的步骤,直至所述训练集达到预设的精度。
- [0034] 可选的,所述处理模块,包括:
- [0035] 特征模块,用于对所述训练图像进行特征处理得到特征区域图像;
- [0036] 确定模块,用于基于所述特征区域图像确定需要进行标注的区域窗口。
- [0037] 本发明实施例还提供了一种电子设备,包括:
- [0038] 存储器和处理器,所述存储器和所述处理器之间互相通信连接,所述存储器中存储有计算机指令,所述处理器通过执行所述计算机指令,从而执行本发明实施例提供的语义分割模型的训练集构建方法。
- [0039] 本发明实施例还提供了一种计算机可读存储介质,所述计算机可读存储介质存储计算机指令,所述计算机指令用于使所述计算机执行本发明实施例提供的语义分割模型的训练集构建方法。
- [0040] 本发明技术方案,具有如下优点:
- [0041] 本发明提供了一种语义分割模型的训练集构建方法,通过获取目标场景的多张训练图像建立未标注图像集合;从未标注图像集合中提取训练图像,并对训练图像进行处理得到训练图像需要进行标注的区域窗口;对区域窗口进行标注并生成标签;将带有标签的区域窗口加入到训练集,并返回从未标注图像集合中提取训练图像的步骤,直至训练集满足预设训练集要求。本发明通过自动筛选不确定性较大的区域并生成区域窗口进行标注,极大地减少了人工标注的人力成本和时间,加快了网络对于特定场景分割任务的训练集构建速度,提高了训练集的构建效率。

附图说明

[0042] 为了更清楚地说明本发明具体实施方式或现有技术中的技术方案,下面将对具体实施方式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,下面描述中的附图是本发明的一些实施方式,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。

[0043] 图1为本发明实施例中的语义分割模型的训练集构建方法的流程图;

[0044] 图2为本发明实施例中的语义分割模型的训练集构建装置的结构示意图;

[0045] 图3为本发明实施例中的电子设备的结构示意图。

具体实施方式

[0046] 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,本领域技术人员在没有做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。

[0047] 根据本发明实施例,提供了一种语义分割模型的训练集构建方法实施例,需要说明的是,在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行,并且,虽然在流程图中示出了逻辑顺序,但是在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤。

[0048] 在本实施例中提供了一种语义分割模型的训练集构建方法,可用于语义分割模型的训练,如图1所示,该语义分割模型的训练集构建方法包括如下步骤:

[0049] 步骤S1:获取目标场景的多张训练图像建立未标注图像集合。具体的,目标场景可以根据不同的客户需求进行选择,直接收集特定的目标场景的RGB图像,而不需要依次通过人工标注它们给定标签,不仅对不同场景的分割任务处理变得更加灵活,同时节省了大量的时间和人力。

[0050] 步骤S2:从未标注图像集合中提取训练图像,并对训练图像进行处理得到训练图像需要进行标注的区域窗口。具体的,通过对训练图像进行特征筛选的处理,使得筛选出的需要标注的区域更加有决定性,间接的再次减少了分割网络达到相同性能下所需的带标签的数据数量。

[0051] 步骤S3:对区域窗口进行标注并生成标签。具体的,由于区域窗口是经过筛选后的一致性较差的图像数据,已将一致性较好的区域筛选了,通过标注界面对筛选出的一致性较差的区域窗口进行人工标注,生成这些区域窗口的标签,大幅度的减少了人工制作语义分割的工作量,节省了时间成本和人力成本。

[0052] 步骤S4:将带有标签的区域窗口加入到训练集,并返回从未标注图像集合中提取训练图像的步骤,直至训练集满足预设训练集要求。具体的,将筛选出的区域窗口的图像和标签加入到训练集中,并将处理后的训练图像从未标注的图像集合中删除。

[0053] 通过上述步骤S1至步骤S4,本发明实施例提供的语义分割模型的训练集构建方法,通过自动筛选不确定性较大的区域并生成区域窗口进行标注,极大地减少了人工标注的人力成本和时间,加快了网络对于特定场景分割任务的训练集构建速度,提高了训练集的构建效率。

[0054] 具体的,在一实施例中,还包括训练语义分割模型,在通过上述步骤构建语义分割模型的训练集的过程中,可以同时进行模型的训练过程,直至语义分割模型性能满足特定任务的要求,停止模型训练;由于语义分割模型的性能达到了要求,说明训练集的构建也达到了要求,可以停止构建训练集。此过程有效提高了模型构建的效率,通过筛选不确定性较大的区域生并进行人工标注生成标签,相较于全图人工标注,极大地减少了人工标注的人力成本和时间。

[0055] 具体地,在一实施例中,上述的步骤S2,具体包括如下步骤:

[0056] 步骤S21:对训练图像进行特征处理得到特征区域图像。具体的,优选采用MobilenetV2轻量级特征提取网络,输入的RGB图像为3通道,经过初始卷积将通道数转为32通道,随后经过17个反向残差块提取单张RGB图像的特征值,有效提高了特征提取精度,进而得到精度较高的特征区域图像。

[0057] 步骤S22:基于特征区域图像确定需要进行标注的区域窗口。具体的,采用区域级数据的筛选方式相比于整张图像的筛选方式,在后续的标注过程中能够大幅减少人工制作数据集时标注的时间成本和工作量。

[0058] 具体地,在一实施例中,上述的步骤S21,具体包括如下步骤:

[0059] 步骤S211:通过预设算法提取训练图像上的多个特征点图像。具体的,通过对特征点图像进行提取,可以有效提高特征识别精度,从而提高训练集构建的效率。

[0060] 步骤S212:将特征点图像分别代入多个预设模型,得到特征点图像的像素分类结果。

[0061] 步骤S213:对像素分类结果的一致性进行判断。

[0062] 步骤S214:选取一致性判断结果未达到预设目标的特征点图像作为特征区域图像。

[0063] 具体的,经过高斯注意卷积、区域推荐算法和非极大值抑制等方式,对像素进行分类,从而选择出一致性较差的数据。一致性判断结果可以有效反映出数据的不确定性分布,在区域的选择过程中,都是以信息熵最高(例如同一个像素点代入10个预设模型,得到10个分类结果,有4个是分类为A,3个分类为B,3个分类为C的话,就代表这个像素点的一致性较差,区域特征不明显,信息熵较高;若10个结果中有8个都认为是A类,2个认为是B,那么一致性较好,区域特征清楚明显,信息熵就较低)的像素点作为区域中心后来做区域的划分选择。相比于普通的卷积,本实施例中的高斯注意卷积更能突出区域中心像素的权重和信息,减少邻域像素对中心像素点的影响,防止引入无关的平滑信息。

[0064] 二维的高斯分布具体表示如下:

$$[0065] \quad G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

[0066] 上式中的函数与正态分布密度函数十分相近,因此采用python自带的一些函数组件来生成一组符合正态分布的数组,之后通过numpy中的outer函数计算矩阵外积并开平方,最终得到一个符合高斯分布的高斯核注意模块。本申请实施例中列举的高斯核注意模块设置为 64×64 的大小,标准差设置为2。通过高斯注意模块来代替普通的卷积核,使得筛选出的区域更加趋向于占比较少或网络学习困难的类别区域,从而平衡数据集中类别不均衡的问题。

[0067] 具体地,在一实施例中,上述的步骤S22,具体包括如下步骤:

[0068] 步骤S221:筛选特征区域图像中信息熵最大的像素点坐标。具体的,输入的特征区域图像经过高斯注意卷积核后得到新的特征图,维度大小取决于模块和筛选窗口的大小。将特征图转成向量的表示方式,并用搜索最大数索引的函数(argmax)进行第一个区域窗口的像素中心筛选。具体公式如下:

$$[0069] \quad P_{\text{index}} = \text{argmax}(\text{Vectorization}(I_M))$$

[0070] 其中, P_{index} 表示整个特征区域图像 I_M 中信息熵值最大的像素点,通过向量化(Vectorization)来将信息图转换成向量的形式,之后经过argmax函数寻找到向量中信息熵最大的像素点。得到最大信息熵的像素点索引之后,通过以下公式得到该像素点的坐标:

$$[0071] \quad P_r = \text{round}(P_{\text{index}}/C_M)$$

$$[0072] \quad P_c = P_{\text{index}} \% C_M$$

[0073] 在上述公式中, P_r 代表熵最大的像素点的横坐标,通过将像素点的索引除以特征图的高度大小得到横坐标 P_r ; P_c 代表像素点的纵坐标,通过对特征图的高度大小进行取余计算获得纵坐标 P_c 。

[0074] 步骤S222:将像素点坐标作为中心点对训练图像进行划分,得到需要进行标注的区域窗口。

[0075] 具体地,在一实施例中,上述的步骤S222,具体包括如下步骤:

[0076] 步骤S2221:获取训练图像的边界信息。

[0077] 步骤S2222:根据像素点坐标和预设的窗口尺寸确定预选区域。

[0078] 步骤S2223:根据边界信息对预选区域进行调整,得到区域窗口。

[0079] 具体的,将筛选出的像素点作为区域窗口的中心点并划分出预选区域,划分窗口的公式如下:

$$[0080] \quad [W_{r1}:W_{r2}] = I_M[\max(0, P_r - W_L) : \min(R_M, P_r + W_L)]$$

$$[0081] \quad [W_{c1}:W_{c2}] = I_M[\max(0, P_c - W_L) : \min(C_M, P_c + W_L)]$$

[0082] 其中, $[W_{r1}:W_{r2}]$ 表示划分区域的横向区域位置, $[W_{c1}:W_{c2}]$ 表示划分区域的纵向位置。 W_L 是划分窗口的大小, R_M 表示信息特征图的横向最大边界, C_M 表示信息特征图的纵向最大边界值。上述公式通过max、min两种函数来确保筛选出的窗口不会超过特征图的边界。被选中的区域会将值都置为0,防止下一轮再次选择,保证区域筛选的有效性和独立性。

[0083] 具体地,在一实施例中,上述的步骤S3,还包括如下步骤:

[0084] 步骤S31:将需要进行标注的区域窗口加入到待处理集合;

[0085] 步骤S32:对待处理集合内的区域窗口进行人工标注,并生成标签。

[0086] 具体的,通过筛选一致性较差的区域窗口并进行人工标注,可以平衡类别信息不对等的情况,以一小部分的标注样本得到较高的分割精度,从而一定程度上减少人工制作数据集的成本。

[0087] 在本实施例中还提供了一种语义分割模型的训练集构建装置,该装置用于实现上述实施例及优选实施方式,已经进行过说明的不再赘述。如以下所使用的,术语“模块”可以实现预定功能的软件和/或硬件的组合。尽管以下实施例所描述的装置较佳地以软件来实现,但是硬件,或者软件和硬件的组合的实现也是可能并被构想的。

[0088] 本实施例提供一种语义分割模型的训练集构建装置,如图2所示,包括:

[0089] 获取模块101,用于获取目标场景的多张训练图像构建未标注图像集合,详细内容参见上述方法实施例中步骤S1的相关描述,在此不再进行赘述。

[0090] 处理模块102,用于从未标注图像集合中提取训练图像,并对训练图像进行处理得到训练图像需要进行标注的区域窗口,详细内容参见上述方法实施例中步骤S2的相关描述,在此不再进行赘述。

[0091] 标注模块103,用于对区域窗口进行标注并生成标签,详细内容参见上述方法实施例中步骤S3的相关描述,在此不再进行赘述。

[0092] 集合模块104,用于将带有标签的区域窗口加入到训练集,并返回从未标注图像集合中提取训练图像的步骤,直至训练集达到预设的精度,详细内容参见上述方法实施例中步骤S4的相关描述,在此不再进行赘述。

[0093] 具体地,在一实施例中,上述的处理模块102,包括:

[0094] 特征模块1021,用于对训练图像进行特征处理得到特征区域图像,详细内容参见上述方法实施例中步骤S21的相关描述,在此不再进行赘述。

[0095] 确定模块1022,用于基于特征区域图像确定需要进行标注的区域窗口,详细内容参见上述方法实施例中步骤S22的相关描述,在此不再进行赘述。

[0096] 本实施例中的语义分割模型的训练集构建装置是以功能单元的形式来呈现,这里的单元是指ASIC电路,执行一个或多个软件或固定程序的处理器和存储器,和/或其他可以提供上述功能的器件。

[0097] 上述各个模块的更进一步的功能描述与上述对应实施例相同,在此不再赘述。

[0098] 根据本发明实施例还提供了一种电子设备,如图3所示,该电子设备可以包括处理器901和存储器902,其中处理器901和存储器902可以通过总线或者其他方式连接,图3中通过总线连接为例。

[0099] 处理器901可以为中央处理器(Central Processing Unit,CPU)。处理器901还可以为其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现场可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等芯片,或者上述各类芯片的组合。

[0100] 存储器902作为一种非暂态计算机可读存储介质,可用于存储非暂态软件程序、非暂态计算机可执行程序以及模块,如本发明方法实施例中的方法所对应的程序指令/模块。处理器901通过运行存储在存储器902中的非暂态软件程序、指令以及模块,从而执行处理器的各种功能应用以及数据处理,即实现上述方法实施例中的方法。

[0101] 存储器902可以包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需要的应用程序;存储数据区可存储处理器901所创建的数据等。此外,存储器902可以包括高速随机存取存储器,还可以包括非暂态存储器,例如至少一个磁盘存储器件、闪存器件、或其他非暂态固态存储器件。在一些实施例中,存储器902可选包括相对于处理器901远程设置的存储器,这些远程存储器可以通过网络连接至处理器901。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0102] 一个或者多个模块存储在存储器902中,当被处理器901执行时,执行上述方法实施例中的方法。

[0103] 上述电子设备具体细节可以对应参阅上述方法实施例中对应的相关描述和效果进行理解,此处不再赘述。

[0104] 本领域技术人员可以理解,实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,的程序可存储于一计算机可读取存储介质中,该程序在执行时,可包括如上述各方法的实施例的流程。其中,存储介质可为磁碟、光盘、只读存储记忆体(Read-Only Memory,ROM)、随机存储记忆体(Random Access Memory,RAM)、快闪存储器(Flash Memory)、硬盘(Hard Disk Drive,缩写:HDD)或固态硬盘(Solid-State Drive,SSD)等;存储介质还可以包括上述种类的存储器的组合。

[0105] 虽然结合附图描述了本发明的实施例,但是本领域技术人员可以在不脱离本发明的精神和范围的情况下做出各种修改和变型,这样的修改和变型均落入由所附权利要求所限定的范围之内。

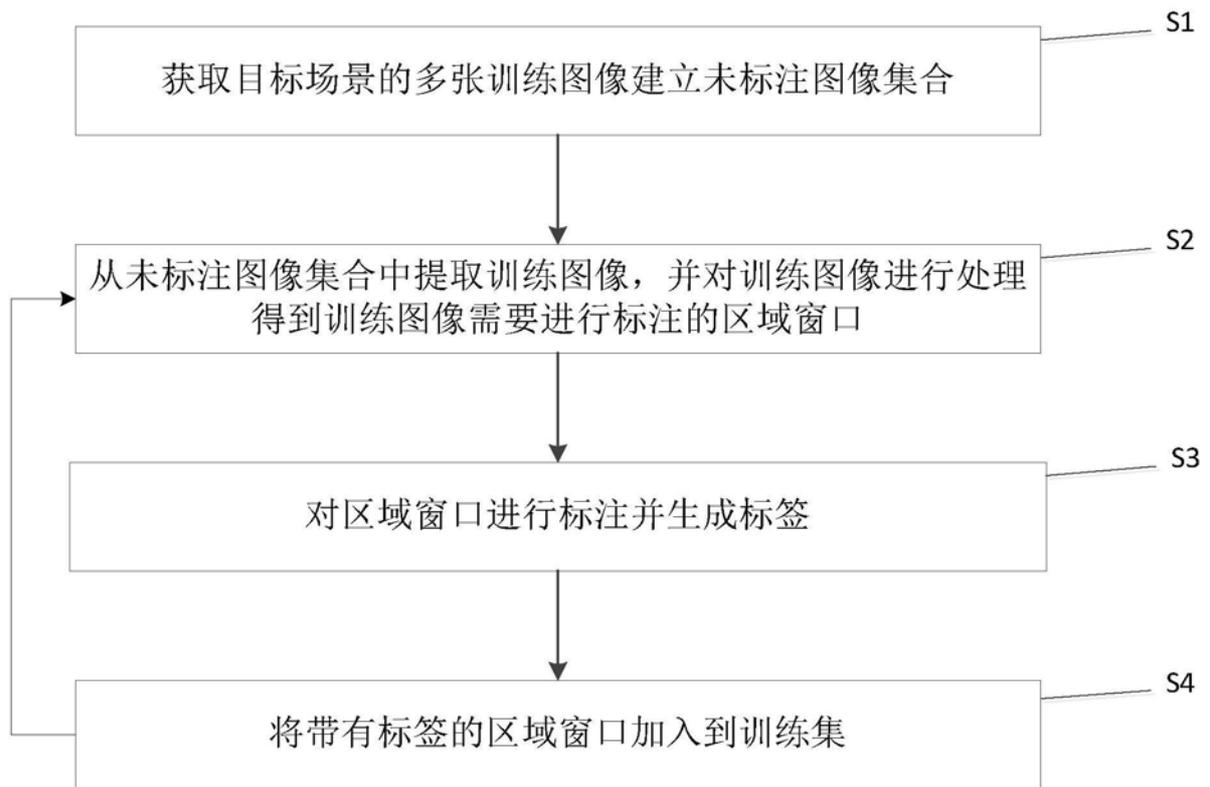


图1



图2

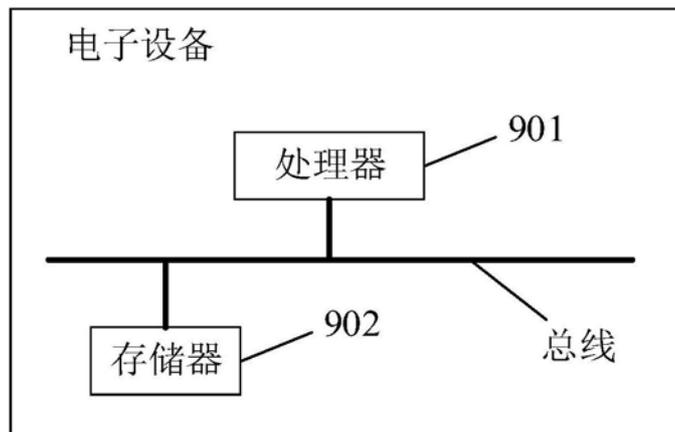


图3