

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2019 年 12 月 19 日 (19.12.2019)



(10) 国际公布号
WO 2019/237342 A1

(51) 国际专利分类号:
G06K 9/00 (2006.01)

北京市朝阳区工体北路甲2号盈科中心捌坊6号3层308单元, Beijing 100027 (CN)。

(21) 国际申请号: PCT/CN2018/091542

(22) 国际申请日: 2018 年 6 月 15 日 (15.06.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 富士通株式会社(FUJITSU LIMITED) [JP/JP]; 日本神奈川县川崎市中原区上小田中4丁目1番1号, Kanagawa 221-8588 (JP)。

(74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街35号国际企业大厦A座16层, Beijing 100033 (CN)。

(72) 发明人; 及

(71) 申请人 (仅对US): 石路(SHI, Lu) [CN/CN]; 中国北京市朝阳区工体北路甲2号盈科中心捌坊6号3层308单元, Beijing 100027 (CN)。 张聪(ZHANG, Cong) [CN/CN]; 中国北京市朝阳区工体北路甲2号盈科中心捌坊6号3层308单元, Beijing 100027 (CN)。 王琪(WANG, Qi) [CN/CN]; 中国

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(54) Title: TRAINING METHOD AND APPARATUS FOR CLASSIFICATION NEURAL NETWORK FOR SEMANTIC SEGMENTATION, AND ELECTRONIC DEVICE

(54) 发明名称: 用于语义分割的分类神经网络的训练方法及装置、电子设备

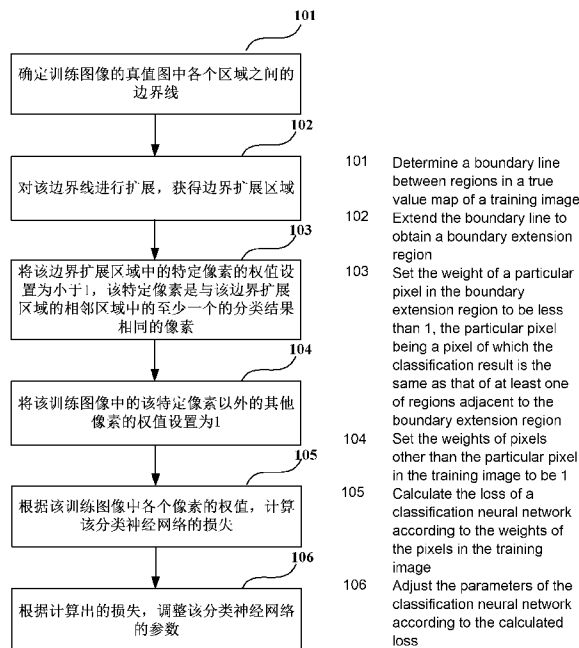


图 1

(57) Abstract: A training method and apparatus for a classification neural network for semantic segmentation, and an electronic device. By setting the weights of uncertain pixels in a boundary region to be small during the calculation of the loss of the neural network, the impact of these pixels on the training speed and model stability can be reduced, thereby effectively increasing the training speed and model stability, and the accuracy of the training result of the main part in a training image can be ensured.

(57) 摘要: 一种用于语义分割的分类神经网络的训练装置及方法、电子设备。通过在计算神经网络的损失时将边界区域中具有不确定性的像素的权值设置为较小, 能够降低这些像素对于训练速度和模型稳定性的影响, 从而有效提高训练速度和模型稳定性, 并且, 能够保证训练图像中主要部分的训练结果的准确性。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

用于语义分割的分类神经网络的训练方法及装置、电子设备

技术领域

本发明涉及信息技术领域，尤其涉及一种语义分割的分类神经网络的训练方法及
5 装置、电子设备。

背景技术

语义分割是结合例如全卷积网络（Fully Convolutional Networks, FCN）等分类神经网络和图像编解码技术的最新技术之一。与其他神经网络操作类似，用于语义分割
10 的数据集非常重要。但与普通神经网络解决的对象分类的课题不同，图像语义分割需要进行的是像素级分割。换句话说，对象检测只需要识别图片的分类，并在此基础上添加位置信息。但是，语义分割需要将所有像素与其位置信息进行分类。它对数据注释有更高的要求。

但是，在数据注释和对于分割效果的要求方面存在如下的问题：某些物体的边界
15 模糊，并且没有准确的边界标记标准；场景的次要部分无法从主要部分分离出来；缺少后处理。目前，出现了一些语义分割算法，针对主要部分能够提供较为准确的分类结果。

应该注意，上面对技术背景的介绍只是为了方便对本发明的技术方案进行清楚、完整的说明，并方便本领域技术人员的理解而阐述的。不能仅仅因为这些方案在本发
20 明的背景技术部分进行了阐述而认为上述技术方案为本领域技术人员所公知。

发明内容

但是，目前的语义分割算法在对细节的处理上存在缺陷，例如，在不同物体的边界区域，存在较大的注释和分类结果的不确定性。并且，由于这些不确定性，导致训练过程中的神经网络损失持续振荡，因此神经网络模型不易收敛，导致训练速度较慢
25 且模型的稳定性较差。

本发明实施例提供一种语义分割的分类神经网络的训练方法及装置、电子设备，通过在计算神经网络的损失时将边界区域中具有不确定性的像素的权值设置为较小，能够降低这些像素对于训练速度和模型稳定性的影响，从而有效提高训练速度和模型

稳定性，并且，能够保证训练图像中主要部分的训练结果的准确性。

根据本发明实施例的第一方面，提供一种用于语义分割的分类神经网络的训练方法，所述方法包括：确定训练图像的真值图中各个区域之间的边界线；对所述边界线进行扩展，获得边界扩展区域；将所述边界扩展区域中的特定像素的权值设置为小于 1，所述特定像素是与所述边界扩展区域的相邻区域中的至少一个的分类结果相同的像素；将所述训练图像中的所述特定像素以外的其他像素的权值设置为 1；根据所述训练图像中各个像素的权值，计算所述分类神经网络的损失；以及根据计算出的所述损失，调整所述分类神经网络的参数。

根据本发明实施例的第二方面，提供一种用于语义分割的分类神经网络的训练装置，所述装置包括：确定单元，其用于确定训练图像的真值图中各个区域之间的边界线；扩展单元，其用于对所述边界线进行扩展，获得边界扩展区域；第一设置单元，其用于将所述边界扩展区域中的特定像素的权值设置为小于 1，所述特定像素是与所述边界扩展区域的相邻区域中的至少一个的分类结果相同的像素；第二设置单元，其用于将所述训练图像中的所述特定像素以外的其他像素的权值设置为 1；计算单元，其用于根据所述训练图像中各个像素的权值，计算所述分类神经网络的损失；以及调整单元，其用于根据计算出的所述损失，调整所述分类神经网络的参数。

根据本发明实施例的第三方面，提供一种电子设备，所述电子设备包括根据本发明实施例的第二方面所述的装置。

本发明的有益效果在于：通过在计算神经网络的损失时将边界区域中具有不确定性的像素的权值设置为较小，能够降低这些像素对于训练速度和模型稳定性的影响，从而有效提高训练速度和模型稳定性，并且，能够保证训练图像中主要部分的训练结果的准确性。

参照后文的说明和附图，详细公开了本发明的特定实施方式，指明了本发明的原理可以被采用的方式。应该理解，本发明的实施方式在范围上并不因而受到限制。在所附权利要求的精神和条款的范围内，本发明的实施方式包括许多改变、修改和等同。

针对一种实施方式描述和/或示出的特征可以以相同或类似的方式在一个或多个其它实施方式中使用，与其它实施方式中的特征相组合，或替代其它实施方式中的特征。

应该强调，术语“包括/包含”在本文使用时指特征、整件、步骤或组件的存在，

但并不排除一个或更多个其它特征、整件、步骤或组件的存在或附加。

附图说明

所包括的附图用来提供对本发明实施例的进一步的理解，其构成了说明书的一部分，用于例示本发明的实施方式，并与文字描述一起来阐释本发明的原理。显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。在附图中：

图 1 是本发明实施例 1 的用于语义分割的分类神经网络的训练方法的一示意图；

图 2 是本发明实施例 1 的训练图像；

10 图 3 是图 2 所示的训练图像的真值图；

图 4 是对图 3 所示的真值图进行线检测后获得的边界线图像；

图 5 是对图 4 所示的边界线图像实施边界线扩展后获得的图像；

图 6 是本发明实施例 1 的用于语义分割的分类神经网络的训练方法的另一示意图；

图 7 是本发明实施例 2 的用于语义分割的分类神经网络的训练装置的一示意图；

15 图 8 是本发明实施例 3 的电子设备的一示意图；

图 9 是本发明实施例 3 的电子设备的系统构成的一示意框图。

具体实施方式

参照附图，通过下面的说明书，本发明的前述以及其它特征将变得明显。在说明书和附图中，具体公开了本发明的特定实施方式，其表明了其中可以采用本发明的原则的部分实施方式，应了解的是，本发明不限于所描述的实施方式，相反，本发明包括落入所附权利要求的范围内的全部修改、变型以及等同物。

实施例 1

本发明实施例提供一种用于语义分割的分类神经网络的训练方法。图 1 是本发明
25 实施例 1 的用于语义分割的分类神经网络的训练方法的一示意图。如图 1 所示，该方法包括：

步骤 101：确定训练图像的真值图中各个区域之间的边界线；

步骤 102：对该边界线进行扩展，获得边界扩展区域；

步骤 103：将该边界扩展区域中的特定像素的权值设置为小于 1，该特定像素是

与该边界扩展区域的相邻区域中的至少一个的分类结果相同的像素；

步骤 104：将该训练图像中的该特定像素以外的其他像素的权值设置为 1；

步骤 105：根据该训练图像中各个像素的权值，计算该分类神经网络的损失；以及

5 步骤 106：根据计算出的损失，调整该分类神经网络的参数。

由上述实施例可知，通过在计算神经网络的损失时将边界区域中具有不确定性的像素的权值设置为较小，能够降低这些像素对于训练速度和模型稳定性的影响，从而有效提高训练速度和模型稳定性，并且，能够保证训练图像中主要部分的训练结果的准确性。

10 在本实施例中，该分类神经网络可以是各种类型的分类神经网络，例如，FCN 等。

在本实施例中，训练图像可以是各种类型的图像，例如，监控视频图像。

在本实施例中，训练图像可以根据各种方式而获得，例如，对监控视频图像进行裁剪而获得多个训练图像。

在步骤 101 中，可以使用现有方法来获得训练图像的真值图并确定真值图中各个
15 区域之间的边界线。例如，通过人工标记的方法获得真值图，并通过线检测的方法获得边界线。

图 2 是本发明实施例 1 的训练图像；图 3 是图 2 所示的训练图像的真值图；图 4 是对图 3 所示的真值图进行线检测后获得的边界线图像。

如图 3 所示，对图 2 所示的训练图像进行人工标记，大致确定了各个主要部分，
20 例如，道路，树木，隔离带等，但是，例如道路和树木的边界区域较为模糊。如图 4 所示，对图 3 所示的真值图进行线检测后，获得了以线条标记的边界线。

在步骤 102 中，对获得的各个边界线进行了扩展，获得了多个边界扩展区域。

图 5 是对图 4 所示的边界线图像实施边界线扩展后获得的图像。如图 5 所示，经过扩展后的边界线变粗，边界扩展区域包含的像素的数量相比于扩展之前增大。这样，
25 便于忽略边界区域中不确定的细节而注重整体的训练。

在本实施例中，对各个边界线扩展可以使用各种方法，例如，将边界线附近一定范围内的像素划分为边界扩展区域内的像素。

在步骤 103 中，将各个边界扩展区域中的特定像素的权值设置为小于 1，该特定像素是与该边界扩展区域的相邻区域中的至少一个的分类结果相同的像素。

例如，对于某个边界扩展区域，其相邻的两个区域为第一相邻区域和第二相邻区域，其中，第一相邻区域的分类结果为“第一分类”，第二相邻区域的分类结果为“第二分类”，此时，将该边界扩展区域中的分类结果是“第一分类”或者“第二分类”的像素作为特定像素。

5 在本实施例中，在确定了该边界扩展区域中的所有特定像素之后，将这些特定像素的权值设置为 1，以用于计算分类神经网络的损失。

这样，这些特定像素实际上是边界扩展区域中分类结果存在不确定性的像素，对于这些不确定的像素，通过减少其计算损失时的权值，能够降低其对神经网络模型收敛的影响，并且，不会影响主要部分的训练结果。

10 在本实施例中，这些像素以及区域的分类结果是通过其真值而确定的，其真值可以根据分类器的输出而确定。例如，根据 softmax 的输出结果而确定。

在本实施例中，可以将特定像素中的分类结果为第二分类且邻近第一相邻区域的像素、以及特定像素中的分类结果为第一分类且邻近第二相邻区域的像素的权值设置为小于 1 的第一权值，将特定像素中的其他像素的权值设置为小于 1 的第二权值，该
15 第一权值大于第二权值。

例如，将特定像素中的分类结果为第二分类且邻近第一相邻区域的像素，以及特定像素中的分类结果为第一分类且邻近第二相邻区域的像素的权值设置为 0.8，将特定像素中的其他像素的权值设置为 0.5。

20 这样，由于分类结果为第二分类且邻近第一相邻区域的像素分类结果为第一分类且邻近第二相邻区域的像素的分类结果是错误的可能性较大，通过提高这些像素的权值，能够进一步提高训练结果的准确性。

在步骤 104 中，将该训练图像中的该特定像素以外的其他像素的权值设置为 1，也就是说，保持这些像素的初始权值。

25 在步骤 105 中，根据该训练图像中各个像素的权值，计算该分类神经网络的损失。计算该分类神经网络的损失可以使用各种方法，例如，根据该训练图像中的各个像素的权值，计算各个像素的交叉熵的累计值，从而获得该分类神经网络的损失。

例如，可以根据以下的公式 (1) 计算该分类神经网络的损失：

$$L(i) = \frac{B \sum_h \sum_w Weight(j) H(j)}{Bhw} \quad (1)$$

其中， $L(i)$ 表示第 i 个训练图像集， B 表示每个训练图像集中的训练图像的数量，

h 表示训练图像的高度， w 表示该训练图像的宽度， $Weight(j)$ 表示该训练图像中第 j 个像素的权值， $H(j)$ 表示该训练图像中第 j 个像素的交叉熵。

在本实施例中，在计算分类神经网络的损失时，还可以考虑摄像角度。例如，根据训练图像中的各个像素的权值以及各个像素的摄像角度，计算各个像素的交叉熵的
5 累计值，从而获得该分类神经网络的损失。

例如，还可以根据以下的公式（2）计算该分类神经网络的损失：

$$L(i) = \frac{B \sum_h \sum_w Weight(j) * Angle(j) H(j)}{Bhw} \quad (2)$$

其中， $L(i)$ 表示第 i 个训练图像集， B 表示每个训练图像集中的训练图像的数量， h 表示训练图像的高度， w 表示该训练图像的宽度， $Weight(j)$ 表示该训练图像中第 j 个像素的权值， $H(j)$ 表示该训练图像中第 j 个像素的交叉熵， $Angle(j)$ 表示该训练图
10 像中第 j 个像素的摄像角度。

在本实施例中，各个像素的摄像角度可以根据现有方法而获得。例如，根据摄像装置出厂时标记的数据而获得。

这样，在计算网络损失时考虑了各个像素的摄像角度的因素，进一步提升了主要
15 被摄体（摄像角度较大）在训练过程中的作用，忽略了次要被摄体（摄像角度较小）在训练过程中的作用，从而能够进一步提高训练的速度以及模型的稳定性。

在步骤 105 中，还可以在计算该分类神经网络的损失时，针对边界扩展区域中的像素，提高属于特定分类的概率。

例如，该特定分类是用户感兴趣的分类。例如，如果用户比较关注车道的完整性，
20 则可以提高边界扩展区域中的像素被判定为车道的概率。

这样，能够保持用户感兴趣的分类或者重视的类别的完整性。

在本实施例中，该方法还可以包括：将不感兴趣区域的像素的权值设置为小于 1，以用于计算该分类神经网络的损失。此时，将特定像素以及不感兴趣区域的像素之外的其他像素的权值设置为 1。

25 这样，通过将用户不感兴趣区域的像素的权值设置为小于 1，能够降低不感兴趣区域对于训练结果的影响，进一步加快训练速度。

在步骤 106 中，根据计算出的损失，调整该分类神经网络的参数。例如，根据损失对于权重参数和偏置参数的偏导数，调整该分类神经网络的各层的权重参数和偏置参数。

图6是本发明实施例1的用于语义分割的分类神经网络的训练方法的另一示意图。

如图6所示，该方法包括：

步骤601：确定训练图像的真值图中各个区域之间的边界线；

步骤602：对该边界线进行扩展，获得边界扩展区域；

5 步骤603：将该边界扩展区域中的特定像素的权值设置为小于1，该特定像素是与该边界扩展区域的相邻区域中的至少一个的分类结果相同的像素；

步骤604：将该训练图像中的该特定像素以外的其他像素的权值设置为1；

步骤605：根据该训练图像中各个像素的权值，计算该分类神经网络的损失；

步骤606：根据计算出的损失，调整该分类神经网络的参数；

10 步骤607：判断该损失是否小于预设阈值，当判断结果为“是”时，进入步骤608，当判断结果为“否”时，进入步骤609；

步骤608：增加该特定像素的权值，并保持其他像素的权值；

步骤609：判断是否满足收敛条件，当判断结果为“是”时，结束训练，当判断结果为“否”时，进入步骤610；

15 步骤610：选择新的训练图像。

在步骤607和步骤608中，判断该损失是否小于预设阈值，该预设阈值可以根据实际需要而设置。当判断为该损失小于预设阈值时，增加所述特定像素的权值以重新计算该损失并进行参数的调整。这样，在分类神经网络模型趋于稳定的情况下，逐渐增加特定像素的权值，能够在模型稳定的情况下进一步提高训练结果的准确性。

20 在步骤608中，增加的权值的数值可以根据实际情况而确定。例如，每次增加原来权值的20%，增加后的权值最大不超过1。

在步骤609中，该收敛条件可以是预先设定的，例如，循环的此次达到了预设值，或者，网络的损失值降低到预设值。

由上述实施例可知，通过在计算神经网络的损失时将边界区域中具有不确定性的
25 像素的权值设置为较小，能够降低这些像素对于训练速度和模型稳定性的影响，从而有效提高训练速度和模型稳定性，并且，能够保证训练图像中主要部分的训练结果的准确性。

实施例2

本发明实施例还提供一种用于语义分割的分类神经网络的训练装置，其对应于实

实施例 1 的训练方法。图 7 是本发明实施例 2 的用于语义分割的分类神经网络的训练装置的一示意图。如图 7 所示，装置 700 包括：

确定单元 701，其用于确定训练图像的真值图中各个区域之间的边界线；

扩展单元 702，其用于对该边界线进行扩展，获得边界扩展区域；

5 第一设置单元 703，其用于将边界扩展区域中的特定像素的权值设置为小于 1，该特定像素是与该边界扩展区域的相邻区域中的至少一个的分类结果相同的像素；

第二设置单元 704，其用于将训练图像中的特定像素以外的其他像素的权值设置为 1；

10 计算单元 705，其用于根据该训练图像中各个像素的权值，计算该分类神经网络的损失；以及

调整单元 706，其用于根据计算出的损失，调整分类神经网络的参数。

在本实施例中，该装置 700 还可以包括：

第三设置单元 707，其用于当该分类神经网络的损失小于预设阈值时，增加特定像素的权值，以重新计算损失并根据该损失进行分类神经网络的参数的调整。

15 在本实施例中，该装置 700 还可以包括：

第四设置单元 708，其用于将不感兴趣区域的像素的权值设置为小于 1，以用于计算分类神经网络的损失。

在本实施例中，上述各个单元的功能的实现可以参见实施例 1 中的训练方法的各个步骤的实施，此处不再赘述。

20 由上述实施例可知，通过在计算神经网络的损失时将边界区域中具有不确定性的像素的权值设置为较小，能够降低这些像素对于训练速度和模型稳定性的影响，从而有效提高训练速度和模型稳定性，并且，能够保证训练图像中主要部分的训练结果的准确性。

实施例 3

25 本发明实施例还提供了一种电子设备，图 8 是本发明实施例 3 的电子设备的示意图。如图 8 所示，电子设备 800 包括用于语义分割的分类神经网络的训练装置 801，其中，用于语义分割的分类神经网络的训练装置 801 的结构和功能与实施例 2 中的记载相同，此处不再赘述。

图 9 是本发明实施例 3 的电子设备的系统构成的一示意框图。如图 9 所示，电子

设备 900 可以包括中央处理器 901 和存储器 902; 存储器 902 耦合到中央处理器 901。该图是示例性的; 还可以使用其它类型的结构, 来补充或代替该结构, 以实现电信功能或其它功能。

如图 9 所示, 该电子设备 900 还可以包括: 输入单元 903、显示器 904、电源 905。

- 5 在一个实施方式中, 实施例 1 所述的用于语义分割的分类神经网络的训练装置的功能可以被集成到中央处理器 901 中。其中, 中央处理器 901 可以被配置为: 确定训练图像的真值图中各个区域之间的边界线; 对所述边界线进行扩展, 获得边界扩展区域; 将所述边界扩展区域中的特定像素的权值设置为小于 1, 所述特定像素是与所述边界扩展区域的相邻区域中的至少一个的分类结果相同的像素; 将所述训练图像中的
- 10 所述特定像素以外的其他像素的权值设置为 1; 根据所述训练图像中各个像素的权值, 计算所述分类神经网络的损失; 以及根据计算出的所述损失, 调整所述分类神经网络的参数。

- 例如, 所述边界扩展区域的相邻区域包括第一相邻区域和第二相邻区域, 所述第一相邻区域的分类结果为第一分类, 所述第二相邻区域的分类结果为第二分类, 所述
- 15 将所述边界扩展区域中的特定像素的权值设置为小于 1, 包括: 将所述特定像素中的分类结果为第二分类且邻近所述第一相邻区域的像素、以及所述特定像素中的分类结果为第一分类且邻近所述第二相邻区域的像素的权值设置为小于 1 的第一权值, 将所述特定像素中的其他像素的权值设置为小于 1 的第二权值, 所述第一权值大于所述第二权值。

- 20 例如, 所述根据所述训练图像中所有像素的权值, 计算所述分类神经网络的损失, 包括: 根据所述训练图像中的各个像素的权值, 计算各个像素的交叉熵的累计值, 从而获得所述分类神经网络的损失。

- 例如, 所述根据所述训练图像中所有像素的权值, 计算所述分类神经网络的损失, 包括: 根据所述训练图像中的各个像素的权值以及各个像素的摄像角度, 计算各个像
- 25 素的交叉熵的累计值, 从而获得所述分类神经网络的损失。

例如, 中央处理器 901 还可以被配置为: 当所述分类神经网络的所述损失小于预设阈值时, 增加所述特定像素的权值, 以重新计算所述损失并根据所述损失进行所述分类神经网络的参数的调整。

例如, 在计算所述分类神经网络的损失时, 针对所述边界扩展区域中的像素, 提

高属于特定分类的概率。

例如，中央处理器 901 还可以被配置为：将不感兴趣区域的像素的权值设置为小于 1，以用于计算所述分类神经网络的损失。

在另一个实施方式中，实施例 1 所述的用于语义分割的分类神经网络的训练装置
5 可以与中央处理器 901 分开配置，例如可以将用于语义分割的分类神经网络的训练装置配置为与中央处理器 901 连接的芯片，通过中央处理器 901 的控制来实现用于语义分割的分类神经网络的训练装置的功能。

在本实施例中电子设备 900 也并不是必须要包括图 9 中所示的所有部件。

如图 9 所示，中央处理器 901 有时也称为控制器或操作控件，可以包括微处理器
10 或其它处理器装置和/或逻辑装置，中央处理器 901 接收输入并控制电子设备 900 的各个部件的操作。

存储器 902，例如可以是缓存器、闪存、硬驱、可移动介质、易失性存储器、非
易失性存储器或其它合适装置中的一种或更多种。并且中央处理器 901 可执行该存储器 902 存储的该程序，以实现信息存储或处理等。其它部件的功能与现有类似，此处
15 不再赘述。电子设备 900 的各部件可以通过专用硬件、固件、软件或其结合来实现，而不偏离本发明的范围。

由上述实施例可知，通过在计算神经网络的损失时将边界区域中具有不确定性的
像素的权值设置为较小，能够降低这些像素对于训练速度和模型稳定性的影响，从而
有效提高训练速度和模型稳定性，并且，能够保证训练图像中主要部分的训练结果的
20 准确性。

本发明实施例还提供一种计算机可读程序，其中当在用于用于语义分割的分类神
经网络的训练装置或电子设备中执行所述程序时，所述程序使得计算机在所述用于语
义分割的分类神经网络的训练装置或电子设备中执行实施例 1 所述的用于语义分割
的分类神经网络的训练方法。

25 本发明实施例还提供一种存储有计算机可读程序的存储介质，其中所述计算机可
读程序使得计算机在用于语义分割的分类神经网络的训练装置或电子设备中执行实
施例 1 所述的用于语义分割的分类神经网络的训练方法。

结合本发明实施例描述的在用于语义分割的分类神经网络的训练装置中进行训
练的方法可直接体现为硬件、由处理器执行的软件模块或二者组合。例如，图 7 中所

示的功能框图中的一个或多个和/或功能框图的一个或多个组合，既可以对应于计算机程序流程的各个软件模块，亦可以对应于各个硬件模块。这些软件模块，可以分别对应于图 1 所示的各个步骤。这些硬件模块例如可利用现场可编程门阵列（FPGA）将这些软件模块固化而实现。

- 5 软件模块可以位于 RAM 存储器、闪存、ROM 存储器、EPROM 存储器、EEPROM 存储器、寄存器、硬盘、移动磁盘、CD-ROM 或者本领域已知的任何其它形式的存储介质。可以将一种存储介质耦接至处理器，从而使处理器能够从该存储介质读取信息，且可向该存储介质写入信息；或者该存储介质可以是处理器的组成部分。处理器和存储介质可以位于 ASIC 中。该软件模块可以存储在移动终端的存储器中，也可以存储在可
- 10 插入移动终端的存储卡中。例如，若设备（例如移动终端）采用的是较大容量的 MEGA-SIM 卡或者大容量的闪存装置，则该软件模块可存储在该 MEGA-SIM 卡或者大容量的闪存装置中。

- 针对图 7 描述的功能框图中的一个或多个和/或功能框图的一个或多个组合，可以实现为用于执行本申请所描述功能的通用处理器、数字信号处理器（DSP）、专用集
- 15 成电路（ASIC）、现场可编程门阵列（FPGA）或其它可编程逻辑器件、分立门或晶体管逻辑器件、分立硬件组件、或者其任意适当组合。针对图 7 描述的功能框图中的一个或多个和/或功能框图的一个或多个组合，还可以实现为计算设备的组合，例如，DSP 和微处理器的组合、多个微处理器、与 DSP 通信结合的一个或多个微处理器或者任何其它这种配置。

- 20 以上结合具体的实施方式对本发明进行了描述，但本领域技术人员应该清楚，这些描述都是示例性的，并不是对本发明保护范围的限制。本领域技术人员可以根据本发明的精神和原理对本发明做出各种变型和修改，这些变型和修改也在本发明的范围内。

权 利 要 求 书

1、一种用于语义分割的分类神经网络的训练方法，所述方法包括：

确定训练图像的真值图中各个区域之间的边界线；

5 对所述边界线进行扩展，获得边界扩展区域；

将所述边界扩展区域中的特定像素的权值设置为小于 1，所述特定像素是与所述边界扩展区域的相邻区域中的至少一个的分类结果相同的像素；

将所述训练图像中的所述特定像素以外的其他像素的权值设置为 1；

根据所述训练图像中各个像素的权值，计算所述分类神经网络的损失；以及

10 根据计算出的所述损失，调整所述分类神经网络的参数。

2、根据权利要求 1 所述的方法，其中，所述边界扩展区域的相邻区域包括第一相邻区域和第二相邻区域，所述第一相邻区域的分类结果为第一分类，所述第二相邻区域的分类结果为第二分类，

所述将所述边界扩展区域中的特定像素的权值设置为小于 1，包括：

15 将所述特定像素中的分类结果为第二分类且邻近所述第一相邻区域的像素、以及所述特定像素中的分类结果为第一分类且邻近所述第二相邻区域的像素的权值设置为小于 1 的第一权值，将所述特定像素中的其他像素的权值设置为小于 1 的第二权值，所述第一权值大于所述第二权值。

3、根据权利要求 1 所述的方法，其中，所述根据所述训练图像中所有像素的权值，计算所述分类神经网络的损失，包括：

20 根据所述训练图像中的各个像素的权值，计算各个像素的交叉熵的累计值，从而获得所述分类神经网络的损失。

4、根据权利要求 3 所述的方法，其中，所述根据所述训练图像中所有像素的权值，计算所述分类神经网络的损失，包括：

25 根据所述训练图像中的各个像素的权值以及各个像素的摄像角度，计算各个像素的交叉熵的累计值，从而获得所述分类神经网络的损失。

5、根据权利要求 1 所述的方法，其中，所述方法还包括：

当所述分类神经网络的所述损失小于预设阈值时，增加所述特定像素的权值，以重新计算所述损失并根据所述损失进行所述分类神经网络的参数的调整。

6、根据权利要求 1 所述的方法，其中，

在计算所述分类神经网络的损失时，针对所述边界扩展区域中的像素，提高属于特定分类的概率。

7、根据权利要求 1 所述的方法，其中，该方法还包括：

5 将不感兴趣区域的像素的权值设置为小于 1，以用于计算所述分类神经网络的损失。

8、一种用于语义分割的分类神经网络的训练装置，所述装置包括：

确定单元，其用于确定训练图像的真值图中各个区域之间的边界线；

扩展单元，其用于对所述边界线进行扩展，获得边界扩展区域；

10 第一设置单元，其用于将所述边界扩展区域中的特定像素的权值设置为小于 1，所述特定像素是与所述边界扩展区域的相邻区域中的至少一个的分类结果相同的像素；

第二设置单元，其用于将所述训练图像中的所述特定像素以外的其他像素的权值设置为 1；

15 计算单元，其用于根据所述训练图像中各个像素的权值，计算所述分类神经网络的损失；以及

调整单元，其用于根据计算出的所述损失，调整所述分类神经网络的参数。

9、根据权利要求 8 所述的装置，其中，所述边界扩展区域的相邻区域包括第一相邻区域和第二相邻区域，所述第一相邻区域的分类结果为第一分类，所述第二相邻区域的分类结果为第二分类，

所述第一设置单元将所述特定像素中的分类结果为第二分类且邻近所述第一相邻区域的像素、以及所述特定像素中的分类结果为第一分类且邻近所述第二相邻区域的像素的权值设置为小于 1 的第一权值，将所述特定像素中的其他像素的权值设置为小于 1 的第二权值，所述第一权值大于所述第二权值。

25 10、根据权利要求 8 所述的装置，其中，

所述计算单元根据所述训练图像中的各个像素的权值，计算各个像素的交叉熵的累计值，从而获得所述分类神经网络的损失。

11、根据权利要求 10 所述的装置，其中，

所述计算单元根据所述训练图像中的各个像素的权值以及各个像素的摄像角度，

计算各个像素的交叉熵的累计值，从而获得所述分类神经网络的损失。

12、根据权利要求 8 所述的装置，其中，所述装置还包括：

第三设置单元，其用于当所述分类神经网络的所述损失小于预设阈值时，增加所述特定像素的权值，以重新计算所述损失并根据所述损失进行所述分类神经网络的参数的调整。

13、根据权利要求 8 所述的装置，其中，

所述计算单元在计算所述分类神经网络的损失时，针对所述边界扩展区域中的像素，提高属于特定分类的概率。

14、根据权利要求 8 所述的装置，其中，所述装置还包括：

10 第四设置单元，其用于将不感兴趣区域的像素的权值设置为小于 1，以用于计算所述分类神经网络的损失。

15、一种电子设备，所述电子设备包括根据权利要求 8-14 中的任一项所述的装置。

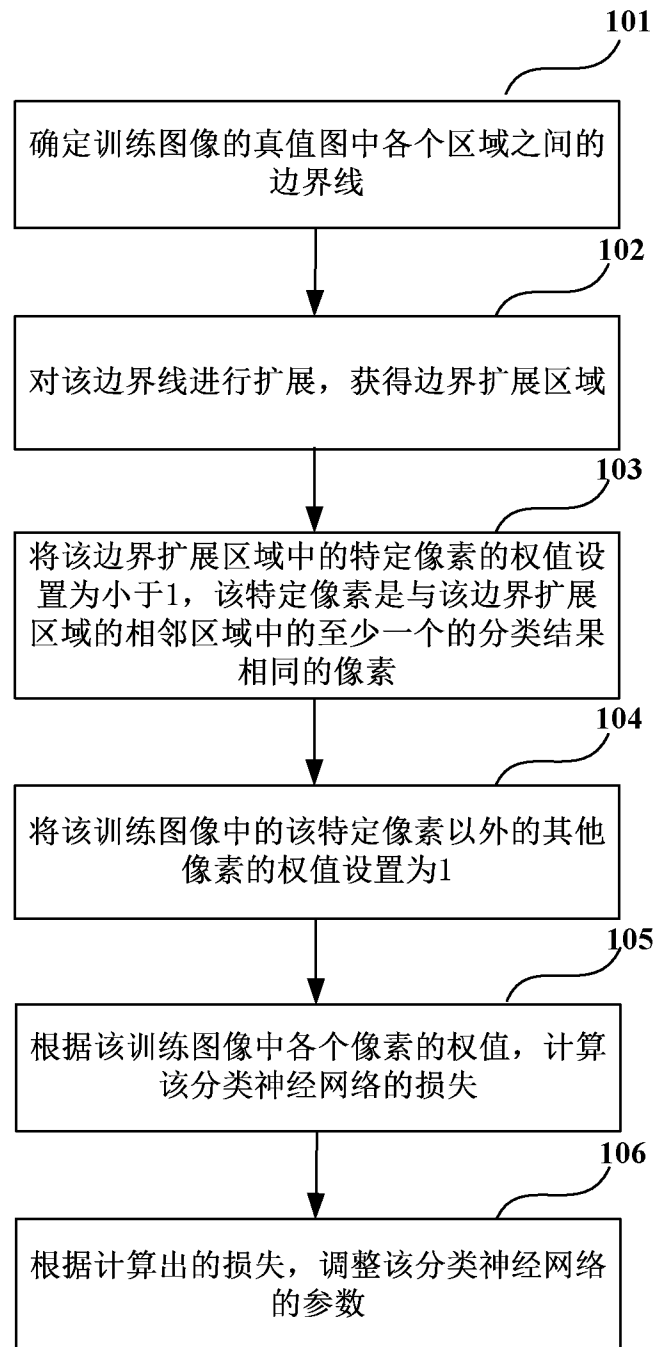


图 1

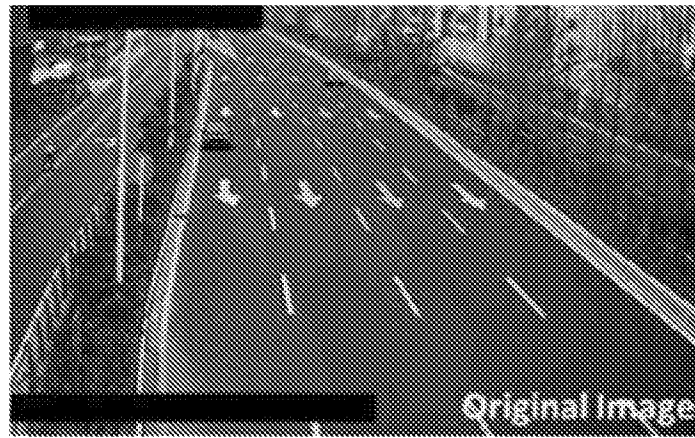


图 2

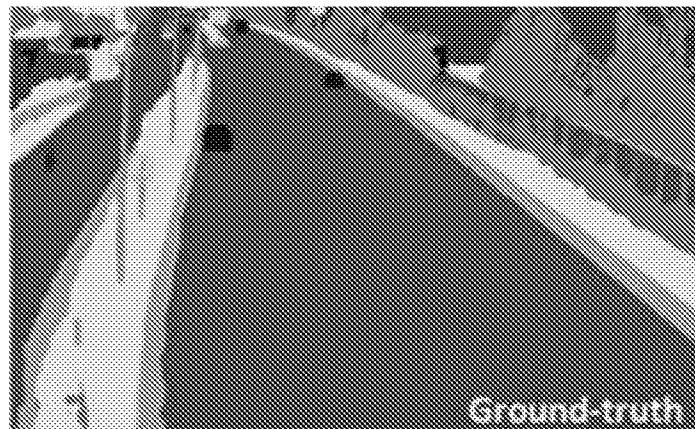


图 3

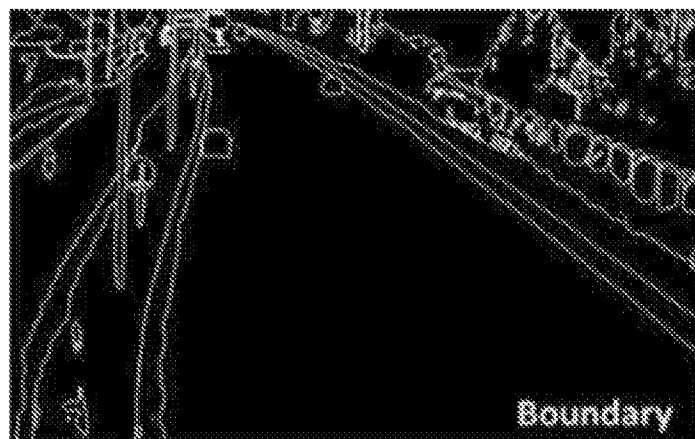


图 4

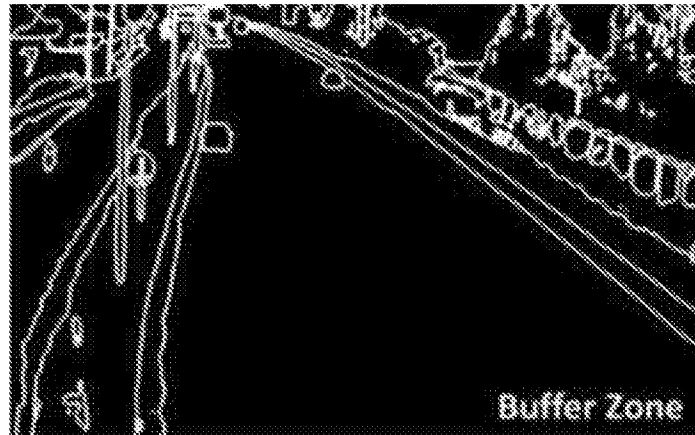


图 5

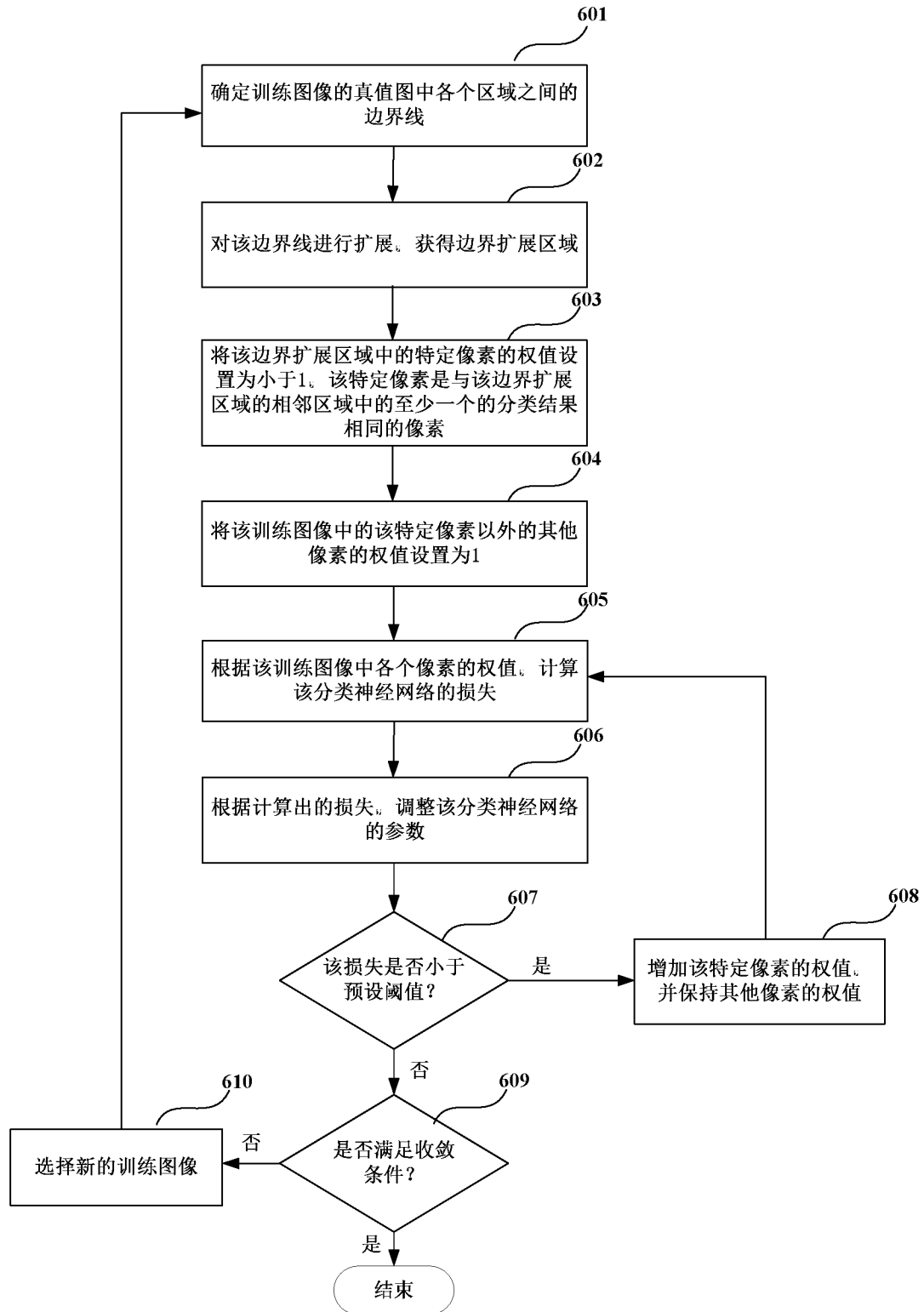


图 6

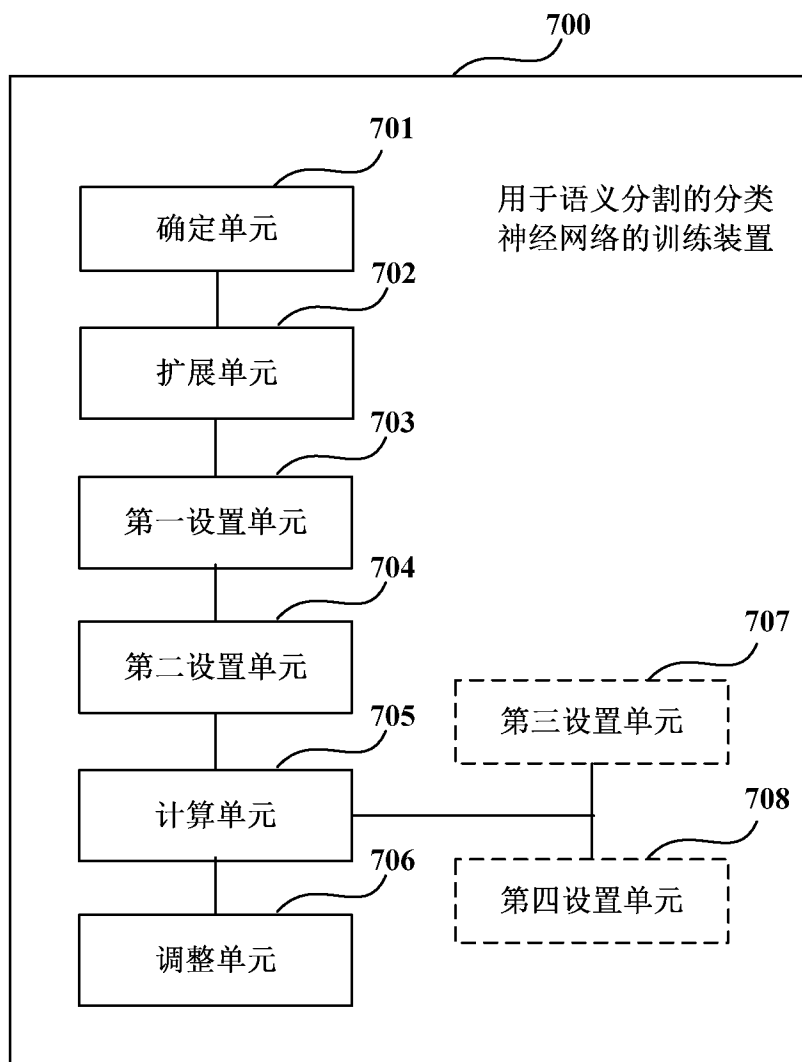


图 7

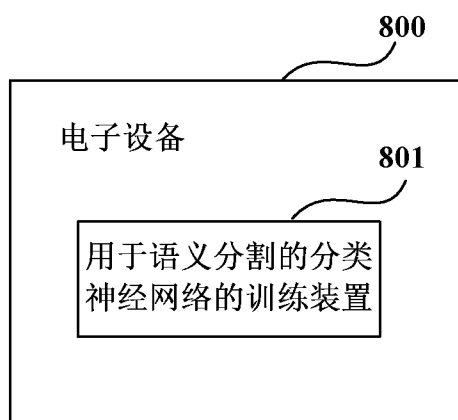


图 8

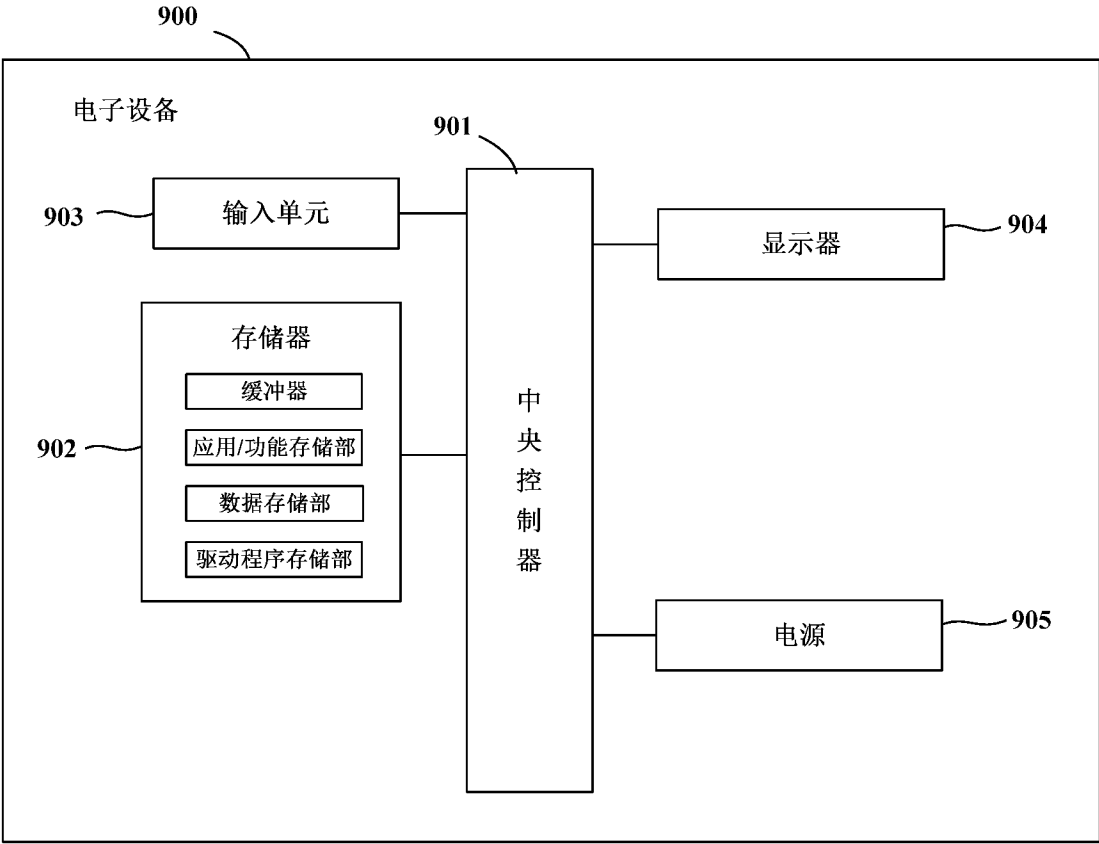


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/091542

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC, IEEE: 语义分割, 分类, 神经网络, 卷积网络, 训练, 损失, 边界, 像素, 权值, semantic, segmentation, classif+, convolution network, neural network, training, loss, boundary, pixel, weight

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 107220980 A (CHONGQING UNIVERSITY OF TECHNOLOGY) 29 September 2017 (2017-09-29) description, paragraphs [0007]-[0017], and figures 2 and 3	1-15
A	CN 107766794 A (TIANJIN UNIVERSITY) 06 March 2018 (2018-03-06) entire document	1-15
A	CN 107945204 A (XIDIAN UNIVERSITY) 20 April 2018 (2018-04-20) entire document	1-15
A	CN 107316307 A (BEIJING UNIVERSITY OF TECHNOLOGY) 03 November 2017 (2017-11-03) entire document	1-15
A	US 2007258648 A1 (XEROX CORPORATION) 08 November 2007 (2007-11-08) entire document	1-15



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

18 February 2019

Date of mailing of the international search report

27 February 2019

Name and mailing address of the ISA/CN

**National Intellectual Property Administration, PRC (ISA/
CN)**
**No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088**
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/091542

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	107220980	A	29 September 2017	None			
CN	107766794	A	06 March 2018	None			
CN	107945204	A	20 April 2018	None			
CN	107316307	A	03 November 2017	None			
US	2007258648	A1	08 November 2007	IN	200700958	I4	28 November 2008

国际检索报告

国际申请号

PCT/CN2018/091542

A. 主题的分类 G06K 9/00 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类		
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06K 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNPAT, CNKI, WPI, EPDOC, IEEE: 语义分割, 分类, 神经网络, 卷积网络, 训练, 损失, 边界, 像素, 权值, semantic, segmentation, classif+, convolution network, neural network, training, loss, boundary, pixel, weight		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 107220980 A (重庆理工大学) 2017年 9月 29日 (2017 - 09 - 29) 说明书第7-17段, 附图2、3	1-15
A	CN 107766794 A (天津大学) 2018年 3月 6日 (2018 - 03 - 06) 全文	1-15
A	CN 107945204 A (西安电子科技大学) 2018年 4月 20日 (2018 - 04 - 20) 全文	1-15
A	CN 107316307 A (北京工业大学) 2017年 11月 3日 (2017 - 11 - 03) 全文	1-15
A	US 2007258648 A1 (XEROX CORPORATION) 2007年 11月 8日 (2007 - 11 - 08) 全文	1-15
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期		国际检索报告邮寄日期
2019年 2月 18日		2019年 2月 27日
ISA/CN的名称和邮寄地址		受权官员
中国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		曹妹妹
		电话号码 86-(10)-53961315

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/091542

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	107220980	A	2017年 9月 29日	无	
CN	107766794	A	2018年 3月 6日	无	
CN	107945204	A	2018年 4月 20日	无	
CN	107316307	A	2017年 11月 3日	无	
US	2007258648	A1	2007年 11月 8日	IN 200700958 I4	2008年 11月 28日