(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
2024 年 1 月 25 日 (25.01.2024)

WIPO | PCT

(10) 国际公布号
WO 2024/017199 A1

(51) 国际专利分类号:
G06V 10/82 (2022.01)

(21) 国际申请号: PCT/CN2023/107725

(22) 国际申请日: 2023 年 7 月 17 日 (17.07.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202210843130.6 2022年7月18日 (18.07.2022) CN

(71) 申请人:北京有竹居网络技术有限公司(**BEIJING YOUZHUJU NETWORK TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市平谷区林荫北街13号信息大厦802室, Beijing 101299 (CN)。

(72) 发明人: 朱敏昭 (**ZHU, Minzhao**); 中国北京市朝阳区七圣中街12号院融中心B1小邮局, Beijing 100028 (CN)。 刘乐遥 (**LIU, Leyao**); 中国北京市朝阳区七圣中街12号院融中心B1小邮局, Beijing 100028 (CN)。 孔涛 (**KONG, Tao**); 中国北京市朝阳区七圣中街12号院融中心B1小邮局, Beijing 100028 (CN)。

(74) 代理人:北京世辉律师事务所 (**SHIHUI PARTNERS**); 中国北京市朝阳区建国门外大街2号北京银泰中心C座42层, Beijing 100022 (CN)。

(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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE,

(54) **Title:** MODEL TRAINING METHOD AND APPARATUS, INSTANCE SEGMENTATION METHOD AND APPARATUS, AND DEVICE AND MEDIUM

(54) 发明名称: 模型训练方法、实例分割方法、装置、设备及介质

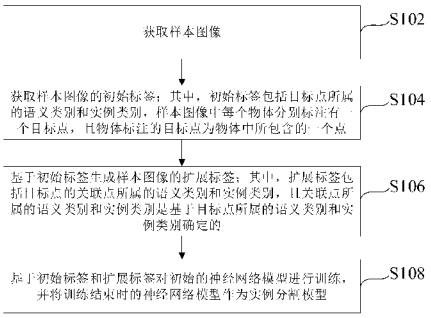


图 1

S102 Acquire a sample image

S104 Acquire an initial label of the sample image, wherein the initial label comprises a semantic category and an instance category to which a target point belongs, each object in the sample image is respectively marked with a target point, and the target point that the object is marked with is a point included in the object

S106 Generate an extension label of the sample image on the basis of the initial label, wherein the extension label comprises a semantic category and an instance category to which an association point of the target point belongs, and the semantic category and the instance category to which the association point belongs are determined on the basis of the semantic category and the instance category to which the target point belongs

S108 Train an initial neural network model on the basis of the initial label and the extension label, and take the neural network model at the end of training as an instance segmentation model

(57) **Abstract:** The embodiments of the present disclosure relate to a model training method and apparatus, an instance segmentation method and apparatus, and a device and a medium. The model training method comprises: acquiring a sample image; acquiring an initial label of the sample image, wherein the initial label comprises a semantic category and an instance category to which a target point belongs, each object in the sample image is respectively marked with a target point, and the target point that the object is marked with is a point included in the object; generating an extension label of the sample image on the basis of the initial label, wherein the extension label comprises a semantic category and an instance category to which an association point of the target point belongs, and the semantic category and the instance category to which the association point belongs are determined on the basis of the semantic category and the instance category to which the target point belongs; and training an initial neural network model on the basis of the initial label and the extension label, and taking the neural network model at the end of training as an instance segmentation model. By means of the embodiments of the present disclosure, a labeling cost required for training a model can be effectively reduced.

WO 2024/017199 A1

[见续页]

BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, 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))。

(57) 摘要: 本公开实施例涉及一种模型训练方法、实例分割方法、装置、设备及介质，其中该模型训练方法包括：获取样本图像；获取样本图像的初始标签；其中，初始标签包括目标点所属的语义类别和实例类别，样本图像中每个物体分别标注有一个目标点，且物体标注的目标点为物体中所包含的一个点；基于初始标签生成样本图像的扩展标签；其中，扩展标签包括目标点的关联点所属的语义类别和实例类别，且关联点所属的语义类别和实例类别是基于目标点所属的语义类别和实例类别确定的；基于初始标签和扩展标签对初始的神经网络模型进行训练，并将训练结束时的神经网络模型作为实例分割模型。本公开实施例可有效节约训练模型所需的标注成本。

模型训练方法、实例分割方法、装置、设备及介质

本申请要求2022年7月18日递交的，标题为“模型训练方法、实例分割方法、装置、设备及介质”、申请号为202210843130.6的中国发明专利申请的优先权。

技术领域

本公开涉及人工智能技术领域，尤其涉及一种模型训练方法、实例分割方法、装置、设备及介质。

背景技术

实例分割任务是计算机视觉中的重要任务，可以预测出图像中每个点所属的语义类别和实例类别，也即可以指出图像中所包含的各种物体类别并进一步区别开不同的实例对象。现有执行实例分割任务的实例分割模型在训练时都需要针对训练样本图像中的大量点进行标注，也即需要非常密集的标注信息，所需标注成本非常高。

发明内容

为了解决上述技术问题或者至少部分地解决上述技术问题，本公开提供了一种模型训练方法、实例分割方法、装置、设备及介质。

本公开实施例提供了一种模型训练方法，包括：获取样本图像；获取所述样本图像的初始标签；其中，所述初始标签包括目标点所属的语义类别和实例类别，所述样本图像中每个物体分别标注有一个目标点，且所述物体标注的目标点为所述物体中所包含的一个点；基于所述初始标签生成所述样本图像的扩展标签；其中，所述扩展标签包括所述目标点的关联点所属的语义类别和实例类别，且所述关联点所属的语义类别和实例类别是基于所述目标点所属的语义类别和实例类别确定的；基于所述初始标签和所述扩展标签对初始的神经网络模型进行训练，并将训练结束时的所述神经网络模型作为实例分割模型。

本公开实施例还提供了一种实例分割方法，包括：获取目标图像；将所述目标图像输入至预先训练得到的实例分割模型；其中，所述实例分割模型是采用上述模型训练方法得到的；获取所述实例分割模型输出的实例分割结果；所述实例分割结果包括所述目标图像中每个点所属的语义类别和实例类别。

本公开实施例还提供了一种模型训练装置，包括：样本获取模块，用于获取样本图像；初始标签获取模块，用于获取所述样本图像的初始标签；其中，所述初始标签包括目标点所属的语义类别和实例类别，所述样本图像中每个物体分别标注有一个目标点，且所述物体标注的目标点为所述物体中所包含的一个点；扩展标签生成模块，用于基于所述初始标签生成所述样本图像的扩展标签；其中，所述扩展标签包括所述目标点的关联点所属的语义类别和实例类别，且所述关联点所属的语义类别和实例类别是基于所述目标点所属的语义类别和实例类别确定的；模型训练模块，用于基于所述初始标签和所述扩展标签对初始的神经网络模型进行训练，并将训练结束时的所述神经网络模型作为实例分割模型。

本公开实施例还提供了一种实例分割装置，包括：图像获取模块，用于获取目标图像；模型输入模块，用于将所述目标图像输入至预先训练得到的实例分割模型；其中，所述实例分割模型是采用任一项所述的模型训练方法得到的；结果获取模块，用于获取所述实例分割模型输出的实例分割结果；所述实例分割结果包括所述目标图像中每个点所属的语义类别和实例类别。

本公开实施例还提供了一种电子设备，所述电子设备包括：处理器；用于存储所述处理器可执行指令的存储器；所述处理器，用于从所述存储器中读取所述可执行指令，并执行所述指令以实现如本公开实施例提供的模型训练方法或者实例分割方法。

本公开实施例还提供了一种计算机可读存储介质，所述存储介质存储有计算机程序，所述计算机程序用于执行如本公开实施例提供的模型训练方法或者实例分割方法。

本公开实施例提供的上述技术方案，在模型训练时可以只针对样本图像中每个物体分别标注一个目标点即可，之后便可基于初始标签（目标点所属的语义类别和实例类别）进行扩展，得到扩展标签（目标点的关联点所属的语义类别和实例类别），从而基于初始标签和扩展标签进行模型训练，便可训练得到实例分割模型及进行后续实例分割。上述方式只需最初针对每个物体标注一个点即可，极大节约了标注成本。且由于实例分割方法所采用的实例分割模型所需的标注成本较低，也即获取实例分割模型更为方便，因此也有助于提升实例分割方法的普适性。

应当理解，本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征，也不用于限制本公开的范围。本公开的其它特征将通过以下的说明书而变得容易理解。

附图说明

此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本公开的实施例，并与说明书一起用于解释本公开的原理。

为了更清楚地说明本公开实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，对于本领域普通技术人员而言，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。

图 1 为本公开实施例提供的一种模型训练方法的流程示意图；

图 2 为本公开实施例提供的一种模型训练示意图；

图 3 为本公开实施例提供的一种实例分割方法的流程示意图；

图 4 为本公开实施例提供的一种模型训练装置的结构示意图；

图 5 为本公开实施例提供的一种实例分割装置的结构示意图；

图 6 为本公开实施例提供的一种电子设备的结构示意图。

具体实施方式

为了能够更清楚地理解本公开的上述目的、特征和优点，下面将对本公开的方案进行进一步描述。需要说明的是，在不冲突的情况下，本公开的实施例及实施例中的特征可以相互组合。

图 1 为本公开实施例提供的一种模型训练方法的流程示意图，该方法可以由模型训练装置执行，其中该装置可以采用软件和/或硬件实现，一般可集成在电子设备中。如图 1 所示，该方法主要包括如下步骤 S102~步骤 S108：

步骤 S102，获取样本图像。

样本图像可以为二维图像（2D 平面图像），也可以为三维图像（诸如三维点云），本公开实施例对样本图像的形式不进行限定。在实际应用中，样本图像的数量通常为多个，且每个样本图像中都会包含至少一种物体，该物体可以为人物、动物、植物、静止的物体或者诸如车辆等动态的物体，应当说明的是，诸如天空、地面等也都可被视为物体，本公开实施例对样本图像中所包含的物体种类/数量不进行限定。

步骤 S104，获取样本图像的初始标签；其中，初始标签包括目标点所属的语义类别和实例类别，样本图像中每个物体分别标注有一个目标点，且物体标注的目标点为物体中所包含的一个点。

初始标签即为最初对样本图像进行标注所得的标签，在本公开实施例中针对样本图像中的每个物体都只需标注一个点即可，物体上标注的点即为目标点，目标点可以为物体中的任意一点，也可以为物体中心点或者物体其它关键位置的点，在此不进行限制。在实际应用中，可以预先标注每个物体上的一个点所属的语义类别和实例类别，从而方便快捷地得到初始标签。上述语义类别可以理解为物体种类，诸如目标点所属的语义类别可以指出目标点所属物体的种类为人物、猫、狗、车辆等；目标点所属的实例类别可以理解为在语义类别的基础上

进一步指出目标点属于具体的哪个物体，诸如在样本图像中包含有多个人时，语义类别仅指出目标点属于人物，而实例类别还可以进一步对多个人物进行区分，指出目标点具体属于哪个人。

在本公开实施例中，只需为每个物体标注一个点的信息即可，假设样本图像中有N个物体，则每个物体标注一个点，最终只需标注N个点（目标点），与相关技术中需要人工在样本图像上标注非常密集的点相比，上述只需针对每个物体标注一个点的方式非常方便快捷，易于实现。

步骤S106，基于初始标签生成样本图像的扩展标签；其中，扩展标签包括目标点的关联点所属的语义类别和实例类别，且关联点所属的语义类别和实例类别是基于目标点所属的语义类别和实例类别确定的。

由于初始只针对每个物体标注一个点，所得的初始标签数量较少，为了能够正常训练模型，因此可以基于初始标签进行扩展，自动生成扩展标签。具体而言，可以找到目标点的关联点，目标点的关联点即为与目标点具有一定关联性的点，诸如，与目标点具有相同特征（诸如颜色等特征）的点和/或位于目标点指定距离范围内的点，已知目标点所属的语义类别和实例类别的基础上，可以进一步扩展获知目标点的关联点所属的语义类别和实例类别，诸如，令关联点所属的语义类别和实例类别与目标点所属的语义类别和实例类别保持一致。通过上述方式可以快速有效地自动扩充标签数量，并在一定程度上保证扩展标签的准确性。

步骤S108，基于初始标签和扩展标签对初始的神经网络模型进行训练，并将训练结束时的神经网络模型作为实例分割模型。

可以理解的是，对神经网络模型进行训练的过程实质为调整神经网络模型的网络参数的过程，直至达到预设条件时（诸如损失函数收敛，神经网络模型可以输出符合预期的实例分割结果）结束训练，得到实例分割模型。本公开实施例对模型结构不进行限制，诸如可以是

卷积神经网络或者其它网络结构等，任何可用于作为实例分割的网络结构均可。在本公开实施例中，最初只需极少量的标注信息，然后在此基础上适当扩展便可进行模型训练，这种训练方式也可称为弱监督训练。在具体训练过程中，还可以基于神经网络模型的输出结果生成伪标签进行自训练，以进一步提升模型训练精度。

本公开实施例提供的上述技术方案，在模型训练时可以只针对样本图像中每个物体分别标注一个目标点即可，之后便可基于初始标签（目标点所属的语义类别和实例类别）进行扩展，得到扩展标签（目标点的关联点所属的语义类别和实例类别），从而基于初始标签和扩展标签进行模型训练，便可训练得到实例分割模型及进行后续实例分割。上述方式只需最初针对每个物体标注一个点即可，极大节约了标注成本。

在一些实施方式中，基于初始标签生成样本图像的扩展标签的步骤，可以参照如下步骤 a~步骤 c 实现：

步骤 a，获取样本图像的低级特征。示例性地，样本图像是二维平面图像时，低级特征诸如可以包括颜色特征；样本图像是三维点云时，低级特征诸如可以包括颜色特征、法向量特征等；此外，低级特征还可以包括诸如纹理特征、形状特征等，在此不进行限制。

步骤 b，根据低级特征对样本图像进行分割处理，基于分割结果得到目标点的关联点。诸如，根据低级特征将样本图像分割为多个局部单元，则将与目标点位于同一局部单元的点作为目标点的关联点。为了能够较为快速准确地找到目标点的关联点，根据低级特征对样本图像进行分割处理，基于分割结果得到目标点的关联点的步骤，包括下述（1）和（2）中至少一项：

（1）在样本图像为二维图像的情况下，根据低级特征对样本图像进行超像素分割处理，基于分割结果得到目标点的关联点。

诸如，可以根据颜色、亮度、纹理等低级特征将二维图像细分为多个局部单元，每个局部单元均可视为一个超像素，超像素由一系列

位置相邻且低级特征相似的像素点组成的小区域，这些小区域大多保留了后续进一步进行图像分割的有效信息，且一般不会破坏图像中物体的边界信息。具体可以采用超像素分割算法对样本图像进行分割处理，从而基于分割结果得到目标点的关联点。在一种具体的实施示例中，可以将样本图像中与目标点同属于一个超像素的点作为目标点的关联点，具体的，超像素分割结果包括多个超像素，每个超像素都可能会包含多个特征相似且位置相邻的点，对于每个目标点而言，可以将与该目标点同属于一个超像素的其余点均作为该目标点的关联点，也可以将与目标点同属于一个超像素的其余点中的部分点（诸如在超像素中与目标点相邻的点）作为目标点的关联点，在此不进行限制。

（2）在样本图像为三维图像的情况下，根据低级特征对样本图像进行超体素分割处理，基于分割结果得到目标点的关联点。

样本图像为三维图像时，可以采用点云数据表示，也即可以根据诸如颜色、法向量等低级特征对点云进行超体素分割，同样可得到点云的多个局部单元，每个局部单元均可视为一个超体素。超体素由点云中一系列位置相邻且低级特征相似的点组成的小区域。具体可以采用超体素分割算法对样本图像进行分割处理，从而基于分割结果得到目标点的关联点。在一种具体的实施示例中，可以将样本图像中与目标点同属于一个超体素的点作为目标点的关联点，具体的，超像素分割结果包括多个超体素，每个超体素都可能会包含多个特征相似且位置相邻的点，对于每个目标点而言，可以将与该目标点同属于一个超体素的其余点均作为该目标点的关联点，也可以将与目标点同属于一个超体素的其余点中的部分点（诸如在超体素中与目标点相邻的点）作为目标点的关联点，在此不进行限制。

步骤 c，将目标点所属的语义类别和实例类别分别作为关联点所属的语义类别和实例类别，得到样本图像的扩展标签。

关联点与目标点的低级特征相似且关联点通常与目标点相邻，因此关联点与目标点属于同一个物体的概率极大，所以可以将目标点所

属的语义类别和实例类别分别作为关联点所属的语义类别和实例类别，通过上述步骤a~步骤c所提供的方式可以在初期得到较为可靠准确的扩展标签，以便于后续采用数量相对较多的标签进行模型的初始训练。

在一些实施方式中，基于初始标签和扩展标签对预设的神经网络模型进行训练，并将训练结束时的神经网络模型作为实例分割模型的步骤，可以参照如下步骤一~步骤三实现:

步骤一，基于初始标签和扩展标签对初始的神经网络模型进行第一训练，将达到第一训练结束条件时的神经网络模型作为第一神经网络模型。其中，第一训练结束条件可以包括以下中的一种或多种：总损失函数值收敛至第一预设阈值，达到第一训练次数阈值、达到第一训练时长阈值。其中，总损失函数值基于标签与模型输出结果确定。

在实际应用中，可以将样本图像输入至神经网络模型中，获取神经网络模型针对样本图像输出的分割结果，基于标签（初始标签和扩展标签）和预设的损失函数对分割结果进行评估，得到总损失函数值。。在一些具体的实施示例中，在训练期间神经网络模型输出的分割结果可以仅采用样本图像中每个点的实例特征值、语义概率和偏移向量等信息表征，其中，语义概率可直接用于语义分割，实例特征值和偏移向量可用于进行实例分割。然后基于标签（初始标签和扩展标签）对分割结果进行评估，诸如，基于标签和判别损失（Discriminative Loss）函数确定实例特征值对应的第一损失值，基于标签和交叉熵损失（Cross Entropy Loss）函数确定语义概率对应的第二损失值，基于L2损失函数确定偏移向量对应的第三损失值，将第一损失值、第二损失值和第三损失值加权求和得到总损失值（也即前述总损失函数值的简称），基于总损失值调整神经网络模型的参数，直至达到第一训练结束条件。此时即可得到初步训练好的第一神经网络模型，第一神经网络模型具一定的实例分割能力。

步骤二，获取第一神经网络模型针对样本图像中每个待测点的输出结果，基于输出结果生成伪标签；伪标签包括样本图像中每个点所

属的语义类别和实例类别。在实际应用中，可以将样本图像中的每个点均作为待测点，也可以将样本图像中的部分点作为待测点，样本图像中的部分点可以是随机确定的多个点，也可以是指定的多个点；诸如可以随机确定样本图像中预设数量个点作为待测点，也可以根据需求而指定分布在样本图像中不同位置的多个点作为待测点。倘若待测点仅是样本图像中的部分点，还可以进一步设定数量阈值，令待测点的数量需要高于预设数量阈值，以此确保模型训练的可靠性。

输出结果包括实例特征值、语义概率和偏移向量；在此基础上，在基于输出结果生成伪标签时，可以基于样本图像中每个待测点的实例特征值、语义概率和偏移向量，以及目标点所属的语义类别和实例类别，采用预设聚类算法生成伪标签。可以理解的是，伪标签并非人为真正标记的标签，而是模型在训练过程中针对无标记数据的预测结果，由于不一定精准，所以可称为伪标签（或伪标记）。但是伪标签可帮助模型学习到无标注数据中隐藏的信息，通过将模型自生成的伪标签再应用于模型训练过程中进行监督，可较为可靠地达到基于自训练的弱监督学习效果。在本公开实施例中，可以基于神经网络模型输出的样本图像中每个待测点的实例特征值、语义概率和偏移向量按照聚类算法进行聚类，根据聚类结果和已经确知的目标点所属的语义类别和实例类别，便可便捷地得到每个待测点所属的语义类别和实例类别，生成的伪标签也相对准确。另外，本公开实施例对聚类算法可不进行限制，包括但不限于 K-means 聚类算法（K 均值聚类）等。

为便于理解，本公开实施例进一步给出了采用预设聚类算法生成伪标签的具体实施方式，示例性地，可以参照如下步骤 A~步骤 B 实现：

步骤 A，将目标点作为预设聚类算法中的初始聚类中心，基于初始聚类中心和样本图像中每个待测点的实例特征值、语义概率和偏移向量进行聚类处理。

聚类中心也可称为种子，初始聚类中心也可称为初始种子，最初可将目标点作为初始聚类中心，然后进行多次迭代聚类，直至聚类中心不再发生变化。具体而言，可以参照如下实现:

在首次聚类处理时，基于样本图像中每个待测点的实例特征值、语义概率和偏移向量，得到每个待测点与初始聚类中心之间的第一相似度（也可称为第一距离），根据第一相似度确定新的聚类中心。具体的，可以根据每个待测点与每个初始聚类中心对应的第一相似度进行类别划分，将每个待测点归到和其相似度最近的聚类中心所属的类，然后再基于初步得到的类簇重新计算每个聚类中心的位置（诸如将每个类内的所有点取平均值等）。在非首次聚类处理时，基于样本图像中每个待测点的实例特征值、语义概率和偏移向量，得到每个待测点与上一次聚类处理所得的聚类中心之间的第二相似度（也可称为第二距离），根据第二相似度确定新的聚类中心（可参见上述确定新的聚类中心的具体实现方式），直至聚类中心不再改变时结束聚类处理。其中，聚类中心的数量等于目标点的数量，也即等于样本图像中的物体数量。上述相似度在聚类算法中也可被称为距离。通过上述多次迭代处理，反复调整聚类中心，最终在聚类中心不再发生改变时停止聚类，便可以得到较为准确的聚类结果。而且由于本公开实施例将每个物体上标注的目标点作为初始聚类中心，不同目标点对应不同物体，这种方式有助于提升聚类效率，在较少的迭代次数内便可找到最终的聚类中心，可有效节约计算开销，较好地避免了最初选取不合适的初始聚类中心而导致较多的计算开销的问题；另外，上述方式在一定程度上也可保障聚类结果的可靠性，最终所得的每个类簇都可较为准确地对应一个物体。

步骤 B，根据聚类结果以及目标点所属的语义类别和实例类别，得到样本图像中每个待测点所属的语义类别和实例类别。

在一些具体的实施示例中，可以根据聚类结果获取目标点的同簇点；可以理解的是，聚类结果包括多个类簇，同簇点为样本图像中与

目标点同属一个类簇的待测点，聚类结果中的类簇与样本图像中的物体一一对应（也即不同类簇对应的物体不同）；换言之，通过上述聚类结果可以找到与每个目标点同属于一个物体的其余待测点，之后便可直接将目标点所属的语义类别和实例类别分别作为同簇点的语义类别和实例类别。可以理解的是，在样本图像中的每个待测点都有其所属的物体，而每个物体都已标注有一个目标点，因此每个待测点都会对应有与其同属一个物体的目标点。对于每个物体而言，通过上述聚类方式得到该物体的类簇，而该物体对应的类簇中目标点的语义类别和实例类别是已知的，其余待测点的语义类别和实例类别理论上与目标点的语义类别和实例类别是一致的，此时便可根据该物体中目标点的语义类别和实例类别直接确定该物体中其余待测点的语义类别和实例类别，通过上述方式，样本图像中的所有待测点的语义类别和实例类别便均可获知。

通过上述步骤 A 和步骤 B，将目标点作为初始聚类中心的方式可有效提升聚类效率，节约聚类所需的计算开销，并保障聚类结果的可靠性，在此基础上通过聚类方式可以较为准确地找到与目标点同属一个物体的其余待测点，并将每个物体中目标点的语义类别和实例类别直接作为该物体中其余待测点的语义类别和实例类别，从而可以高效且准确地对每个未标注的待测点进行自动标注，得到模型输出的伪标签。

步骤三，基于伪标签对第一神经网络模型进行第二训练，并将达到第二训练结束条件时的第一神经网络模型作为实例分割模型。

基于伪标签对第一神经网络模型进行第二训练的过程与前述基于初始标签和扩展标签对初始的神经网络模型进行第一训练的过程类似，区别在于第一训练为初始训练，采用的是初始标签和基于初始标签扩展得到的标签，第二训练为模型自训练，采用的是初步训练好的第一神经网络模型输出结果所得到的伪标签，也即训练时所采用的标签不同。上述第二训练结束条件可以包括以下中的一种或多种：总损失函

数值收敛至第二预设阈值，达到第二训练次数阈值、达到第二训练时长阈值。上述总损失函数值的计算方式也参照第一训练过程中计算总损失函数值的方式，诸如需要实例特征值对应的第一损失值、语义概率对应的第二损失值和偏移向量对应的第三损失值进行加权求和得到总损失函数值，在此不再赘述。另外应当说明的是，第一训练和第二训练都表示训练阶段，而不代表训练次数，无论是第一训练还是第二训练，都会训练多次（也即通过反向传播的方式多次调整模型参数），每次训练都会输出伪标签，输出的伪标签用以监督模型在下一次训练时的输出结果，以此来调整模型参数，直至达到预设的训练结束条件时停止训练。

通过上述步骤一~步骤三，首先采用初始标签和扩展标签对神经网络模型进行初步训练，在神经网络模型具有一定实例分割能力后，再基于神经网络模型的输出结果生成伪标签，采用伪标签对神经网络模型进行自训练，上述方式最初只需要针对每个物体标注一个点，后续标签均可基于初始标注的点进行自动生成，极大降低了标注成本，而且上述无论是基于低级特征获得扩展标签还是基于聚类方式获得伪标签的方式较为可靠，可以有效保障用于训练模型的标签的准确性，也在一定程度上保障了训练所得的模型精度。

为了便于理解，可参照图 2 所示的一种模型训练示意图，在该示意图中重点示意出将样本图像输入至神经网络模型，神经网络模型可针对样本图像中的每个点分别输出实例特征值、语义概率和偏移向量，基于输出结果便可生成伪标签，伪标签包含样本图像中每个点所属的语义类别和实例类别，其中，语义类别基于语义概率得到，实例类别基于实例特征值和偏移向量得到；然后便可采用伪标签对神经网络模型进行监督，从而调整神经网络模型的参数。在图 2 中示意出每个物体都只需标注一个目标点，同时示意出基于模型的输出结果生成伪标签的方式，具体的，可基于模型针对样本图像输出的每个点的实例特征值、语义概率和偏移向量采用聚类算法进行聚类处理，其中，目标

点为初始聚类中心，分别为O1、O2和O3（以不同灰度的实心圈表示），以随意一个点为例，该点与O1之间的相似度为0.8、与O2之间的相似度为0.6，与O3之间的相似度为0.2，通过计算每个点分别与O1、O2和O3之间的相似度后可以重新聚类，不同类簇对应的灰度不同；然后找到新的类簇的聚类中心O1'、O2'、O3'（以不同灰度的实心圈表示），之后再次计算每个点与新的聚类中心O1'、O2'、O3'之间的相似度，再基于相似度重新聚类，经过多次聚类后最终得到聚类中心不再改变的类簇，每个类簇对应一个物体，从而基于每个点所属类簇中的目标点的语义类别和实例类别，得到每个点的语义类别和实例类别，较为便捷可靠地生成伪标签。另外应当说明的是，图2仅为简单示例，并未将所有类簇/所有点均示意出。

在前述基础上，本公开实施例提供了一种实例分割方法，参见图3所示的一种实例分割方法的流程示意图，主要包括如下步骤S302~步骤S306：

步骤S302，获取目标图像。目标图像即为待进行实例分割的图像。

步骤S304，将目标图像输入至预先训练得到的实例分割模型；其中，实例分割模型是采用上述任一项模型训练方法得到的，具体训练方式可以参照上述相关内容，在此不进行限定。

步骤S306，获取实例分割模型输出的实例分割结果；实例分割结果包括目标图像中每个点所属的语义类别和实例类别。具体实现时，实例分割模型可以针对目标图像输出每个点对应的实例特征值、语义概率和偏移向量，基于语义概率即可得到每个点所属的语义类别，基于实例特征值和语义概率即可计算不同点之间的相似度，从而通过诸如Mean-shift聚类方式得到每个点所属的实例类别。

由于实例分割方法所采用的实例分割模型所需的标注成本较低，也即获取实例分割模型更为方便，因此也有助于提升实例分割方法的普适性。进一步，基于前述方式无论是基于低级特征获得扩展标签还是基于聚类方式获得伪标签的方式均较为可靠，相比于常规的弱监督

训练方式而言，本公开实施例提供的上述方式可以有效保障用于训练模型的标签的准确性，因此训练所得的实例分割模型的精确度较高，也即本公开实施例只需极少的标注信息便可达到较高的分割精度。

对应于前述模型训练方法，本公开实施例提供了一种模型训练装置的结构示意图，图 4 为本公开实施例提供的一种模型训练装置的结构示意图，该装置可由软件和/或硬件实现，一般可集成在电子设备中，如图 4 所示，包括：

样本获取模块 402，用于获取样本图像；

初始标签获取模块 404，用于获取样本图像的初始标签；其中，初始标签包括目标点所属的语义类别和实例类别，样本图像中每个物体分别标注有一个目标点，且物体标注的目标点为物体中所包含的一个点；

扩展标签生成模块 406，用于基于初始标签生成样本图像的扩展标签；其中，扩展标签包括目标点的关联点所属的语义类别和实例类别，且关联点所属的语义类别和实例类别是基于目标点所属的语义类别和实例类别确定的；

模型训练模块 408，用于基于初始标签和扩展标签对初始的神经网络模型进行训练，并将训练结束时的神经网络模型作为实例分割模型。

本公开实施例提供的上述技术方案，在模型训练时可以只针对样本图像中每个物体分别标注一个目标点即可，之后便可基于初始标签（目标点所属的语义类别和实例类别）进行扩展，得到扩展标签（目标点的关联点所属的语义类别和实例类别），从而基于初始标签和扩展标签进行模型训练，便可训练得到实例分割模型及进行后续实例分割。上述方式只需最初针对每个物体标注一个点即可，极大节约了标注成本。

在一些实施方式中，扩展标签生成模块 406 具体用于：获取所述样本图像的低级特征；根据所述低级特征对所述样本图像进行分割处理，基于分割结果得到所述目标点的关联点；将所述目标点所属的语

义类别和实例类别分别作为所述关联点所属的语义类别和实例类别，得到所述样本图像的扩展标签。

在一些实施方式中，扩展标签生成模块406具体用于下述至少一项：在所述样本图像为二维图像的情况下，根据所述低级特征对所述样本图像进行超像素分割处理，基于分割结果得到所述目标点的关联点；在所述样本图像为三维图像的情况下，根据所述低级特征对所述样本图像进行超体素分割处理，基于分割结果得到所述目标点的关联点。

在一些实施方式中，模型训练模块408具体用于：基于所述初始标签和所述扩展标签对初始的神经网络模型进行第一训练，将达到第一训练结束条件时的神经网络模型作为第一神经网络模型；获取所述第一神经网络模型针对所述样本图像中每个待测点的输出结果，基于所述输出结果生成伪标签；所述伪标签包括所述样本图像中每个所述待测点所属的语义类别和实例类别；基于所述伪标签对所述第一神经网络模型进行第二训练，并将达到第二训练结束条件时的第一神经网络模型作为实例分割模型。

在一些实施方式中，所述输出结果包括实例特征值、语义概率和偏移向量；模型训练模块408具体用于：基于所述样本图像中每个所述待测点的实例特征值、语义概率和偏移向量，以及所述目标点所属的语义类别和实例类别，采用预设聚类算法生成伪标签。

在一些实施方式中，模型训练模块408具体用于：将所述目标点作为预设聚类算法中的初始聚类中心，基于所述初始聚类中心和所述样本图像中每个所述待测点的实例特征值、语义概率和偏移向量进行聚类处理；根据聚类结果以及所述目标点所属的语义类别和实例类别，得到所述样本图像中每个所述待测点所属的语义类别和实例类别。

在一些实施方式中，模型训练模块408具体用于：在首次聚类处理时，基于所述样本图像中每个所述待测点的实例特征值、语义概率和偏移向量，得到每个所述待测点与所述初始聚类中心之间的第一相

似度，根据所述第一相似度确定新的聚类中心；在非首次聚类处理时，基于所述样本图像中每个所述待测点的实例特征值、语义概率和偏移向量，得到每个所述待测点与上一次聚类处理所得的聚类中心之间的第二相似度，根据所述第二相似度确定新的聚类中心，直至聚类中心不再改变时结束聚类处理。

在一些实施方式中，模型训练模块408具体用于：根据聚类结果获取所述目标点的同簇点；其中，所述同簇点为所述样本图像中与所述目标点同属一个类簇的所述待测点，所述聚类结果中的类簇与所述样本图像中的物体一一对应；将所述目标点所属的语义类别和实例类别分别作为所述同簇点的语义类别和实例类别。

本公开实施例所提供的模型训练装置可执行本公开任意实施例所提供的模型训练方法，具备执行方法相应的功能模块和有益效果。

对应于前述实例分割方法，图5为本公开实施例提供的一种实例分割装置的结构示意图，该装置可由软件和/或硬件实现，一般可集成在电子设备中。如图5所示，包括：

图像获取模块502，用于获取目标图像；

模型输入模块504，用于将所述目标图像输入至预先训练得到的实例分割模型；其中，所述实例分割模型是采用任一项所述的模型训练方法得到的；

结果获取模块506，用于获取所述实例分割模型输出的实例分割结果；所述实例分割结果包括所述目标图像中每个点所属的语义类别和实例类别。

由于实例分割装置所采用的实例分割模型所需的标注成本较低，也即获取实例分割模型更为方便，因此也有助于提升实例分割方法的普适性。

本公开实施例所提供的实例分割装置可执行本公开任意实施例所提供的实例分割方法，具备执行方法相应的功能模块和有益效果。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的装置实施例的具体工作过程，可以参考方法实施例中的对应过程，在此不再赘述。

本公开实施例提供一种电子设备，电子设备包括：处理器；用于存储处理器可执行指令的存储器；处理器，用于从存储器中读取可执行指令，并执行指令以实现上述任一模型训练方法或者实例分割方法。

图 6 为本公开实施例提供的一种电子设备的结构示意图。如图 6 所示，电子设备 600 包括一个或多个处理器 601 和存储器 602。

处理器 601 可以是中央处理单元（CPU）或者具有数据处理能力和/或指令执行能力的其他形式的处理单元，并且可以控制电子设备 600 中的其他组件以执行期望的功能。

存储器 602 可以包括一个或多个计算机程序产品，所述计算机程序产品可以包括各种形式的计算机可读存储介质，例如易失性存储器和/或非易失性存储器。所述易失性存储器例如可以包括随机存取存储器（RAM）和/或高速缓冲存储器（cache）等。所述非易失性存储器例如可以包括只读存储器（ROM）、硬盘、闪存等。在所述计算机可读存储介质上可以存储一个或多个计算机程序指令，处理器 601 可以运行所述程序指令，以实现上文所述的本公开的实施例的模型训练方法或实例分割方法以及/或者其他期望的功能。在所述计算机可读存储介质中还可以存储诸如输入信号、信号分量、噪声分量等各种内容。

在一个示例中，电子设备 600 还可以包括：输入装置 603 和输出装置 604，这些组件通过总线系统和/或其他形式的连接机构（未示出）互连。

此外，该输入装置 603 还可以包括例如键盘、鼠标等等。

该输出装置 604 可以向外部输出各种信息，包括确定出的距离信息、方向信息等。该输出装置 604 可以包括例如显示器、扬声器、打印机、以及通信网络及其所连接的远程输出设备等等。

当然，为了简化，图 6 中仅示出了该电子设备 600 中与本公开有关的组件中的一些，省略了诸如总线、输入/输出接口等等的组件。除此之外，根据具体应用情况，电子设备 600 还可以包括任何其他适当的组件。

除了上述方法和设备以外，本公开的实施例还可以是计算机程序产品，其包括计算机程序指令，所述计算机程序指令在被处理器运行时使得所述处理器执行本公开实施例所提供的模型训练方法或实例分割方法。

所述计算机程序产品可以以一种或多种程序设计语言的任意组合来编写用于执行本公开实施例操作的程序代码，所述程序设计语言包括面向对象的程序设计语言，诸如 Java、C++等，还包括常规的过程式程序设计语言，诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、作为一个独立的软件包执行、部分在用户计算设备上部分在远程计算设备上执行、或者完全在远程计算设备或服务器上执行。

此外，本公开的实施例还可以是计算机可读存储介质，其上存储有计算机程序指令，所述计算机程序指令在被处理器运行时使得所述处理器执行本公开实施例所提供的模型训练方法或实例分割方法。

所述计算机可读存储介质可以采用一个或多个可读介质的任意组合。可读介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以包括但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。可读存储介质的更具体的例子（非穷举的列表）包括：具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、光纤、便携式紧凑盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。

本公开实施例还提供了一种计算机程序产品，包括计算机程序/指令，该计算机程序/指令被处理器执行时实现本公开实施例中的模型训练方法或实例分割方法。

需要说明的是，在本文中，诸如“第一”和“第二”等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

以上所述仅是本公开的具体实施方式，使本领域技术人员能够理解或实现本公开。对这些实施例的多种修改对本领域的技术人员来说将是显而易见的，本文中所定义的一般原理可以在不脱离本公开的精神或范围的情况下，在其它实施例中实现。因此，本公开将不会被限制于本文所述的这些实施例，而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。

权 利 要 求 书

1．一种模型训练方法，其特征在于，包括：

获取样本图像；

获取所述样本图像的初始标签；其中，所述初始标签包括目标点所属的语义类别和实例类别，所述样本图像中每个物体分别标注有一个目标点，且所述物体标注的目标点为所述物体中所包含的一个点；

基于所述初始标签生成所述样本图像的扩展标签；其中，所述扩展标签包括所述目标点的关联点所属的语义类别和实例类别，且所述关联点所属的语义类别和实例类别是基于所述目标点所属的语义类别和实例类别确定的；

基于所述初始标签和所述扩展标签对初始的神经网络模型进行训练，并将训练结束时的所述神经网络模型作为实例分割模型。

2．根据权利要求1所述的方法，其特征在于，基于所述初始标签生成所述样本图像的扩展标签的步骤，包括：

获取所述样本图像的低级特征；

根据所述低级特征对所述样本图像进行分割处理，基于分割结果得到所述目标点的关联点；

将所述目标点所属的语义类别和实例类别分别作为所述关联点所属的语义类别和实例类别，得到所述样本图像的扩展标签。

3．根据权利要求2所述的方法，其特征在于，根据所述低级特征对所述样本图像进行分割处理，基于分割结果得到所述目标点的关联点的步骤，包括下述至少一项：

在所述样本图像为二维图像的情况下，根据所述低级特征对所述样本图像进行超像素分割处理，基于分割结果得到所述目标点的关联点；

在所述样本图像为三维图像的情况下，根据所述低级特征对所述样本图像进行超体素分割处理，基于分割结果得到所述目标点的关联点。

4．根据权利要求 1-3 任一项所述的方法，其特征在于，基于所述初始标签和所述扩展标签对预设的神经网络模型进行训练，并将训练结束时的所述神经网络模型作为实例分割模型的步骤，包括：

基于所述初始标签和所述扩展标签对初始的神经网络模型进行第一训练，将达到第一训练结束条件时的神经网络模型作为第一神经网络模型；

获取所述第一神经网络模型针对所述样本图像中每个待测点的输出结果，基于所述输出结果生成伪标签；所述伪标签包括所述样本图像中每个所述待测点所属的语义类别和实例类别；

基于所述伪标签对所述第一神经网络模型进行第二训练，并将达到第二训练结束条件时的第一神经网络模型作为实例分割模型。

5．根据权利要求 4 所述的方法，其特征在于，所述输出结果包括实例特征值、语义概率和偏移向量；基于所述输出结果生成伪标签的步骤，包括：

基于所述样本图像中每个所述待测点的实例特征值、语义概率和偏移向量，以及所述目标点所属的语义类别和实例类别，采用预设聚类算法生成伪标签。

6．根据权利要求 5 所述的方法，其特征在于，采用预设聚类算法生成伪标签的步骤，包括：

将所述目标点作为预设聚类算法中的初始聚类中心，基于所述初始聚类中心和所述样本图像中每个所述待测点的实例特征值、语义概率和偏移向量进行聚类处理；

根据聚类结果以及所述目标点所属的语义类别和实例类别，得到所述样本图像中每个所述待测点所属的语义类别和实例类别。

7. 根据权利要求6所述的方法，其特征在于，基于所述初始聚类中心和所述样本图像中每个所述待测点的实例特征值、语义概率和偏移向量进行聚类处理的步骤，包括：

在首次聚类处理时，基于所述样本图像中每个所述待测点的实例特征值、语义概率和偏移向量，得到每个所述待测点与所述初始聚类中心之间的第一相似度，根据所述第一相似度确定新的聚类中心；

在非首次聚类处理时，基于所述样本图像中每个所述待测点的实例特征值、语义概率和偏移向量，得到每个所述待测点与上一次聚类处理所得的聚类中心之间的第二相似度，根据所述第二相似度确定新的聚类中心，直至聚类中心不再改变时结束聚类处理。

8. 根据权利要求6所述的方法，其特征在于，根据聚类结果以及所述目标点所属的语义类别和实例类别，得到所述样本图像中每个所述待测点所属的语义类别和实例类别的步骤，包括：

根据聚类结果获取所述目标点的同簇点；其中，所述同簇点为所述样本图像中与所述目标点同属一个类簇的所述待测点，所述聚类结果中的类簇与所述样本图像中的物体一一对应；

将所述目标点所属的语义类别和实例类别分别作为所述同簇点的语义类别和实例类别。

9. 一种实例分割方法，其特征在于，包括：

获取目标图像；

将所述目标图像输入至预先训练得到的实例分割模型；其中，所述实例分割模型是采用权利要求1至8任一项所述的模型训练方法得到的；

获取所述实例分割模型输出的实例分割结果；所述实例分割结果包括所述目标图像中每个点所属的语义类别和实例类别。

10. 一种模型训练装置，其特征在于，包括：

样本获取模块，用于获取样本图像；

初始标签获取模块，用于获取所述样本图像的初始标签；其中，所述初始标签包括目标点所属的语义类别和实例类别，所述样本图像中每个物体分别标注有一个目标点，且所述物体标注的目标点为所述物体中所包含的一个点；

扩展标签生成模块，用于基于所述初始标签生成所述样本图像的扩展标签；其中，所述扩展标签包括所述目标点的关联点所属的语义类别和实例类别，且所述关联点所属的语义类别和实例类别是基于所述目标点所属的语义类别和实例类别确定的；

模型训练模块，用于基于所述初始标签和所述扩展标签对初始的神经网络模型进行训练，并将训练结束时的所述神经网络模型作为实例分割模型。

11．一种实例分割装置，其特征在于，包括：

图像获取模块，用于获取目标图像；

模型输入模块，用于将所述目标图像输入至预先训练得到的实例分割模型；其中，所述实例分割模型是采用任一项所述的模型训练方法得到的；

结果获取模块，用于获取所述实例分割模型输出的实例分割结果；所述实例分割结果包括所述目标图像中每个点所属的语义类别和实例类别。

12．一种电子设备，其特征在于，所述电子设备包括：

处理器；

用于存储所述处理器可执行指令的存储器；

所述处理器，用于从所述存储器中读取所述可执行指令，并执行所述指令以实现上述权利要求 1-8 中任一所述的模型训练方法或者权利要求 9 所述的实例分割方法。

13．一种计算机可读存储介质，其特征在于，所述存储介质存储有计算机程序，所述计算机程序用于执行上述权利要求 1-8 中任一所述的模型训练方法或者权利要求 9 所述的实例分割方法。

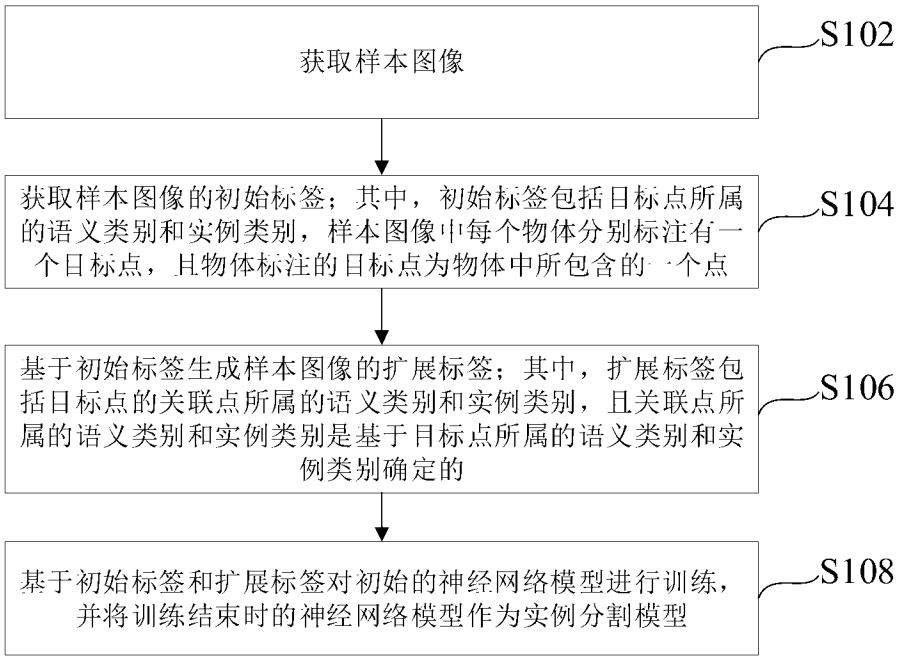

获取样本图像
S102
获取样本图像的初始标签；其中，初始标签包括目标点所属的语义类别和实例类别，样本图像中每个物体分别标注有一个目标点，且物体标注的目标点为物体中所包含的一个点
S104
基于初始标签生成样本图像的扩展标签；其中，扩展标签包括目标点的关联点所属的语义类别和实例类别，且关联点所属的语义类别和实例类别是基于目标点所属的语义类别和实例类别确定的
S106
基于初始标签和扩展标签对初始的神经网络模型进行训练，并将训练结束时的神经网络模型作为实例分割模型
S108

图 1

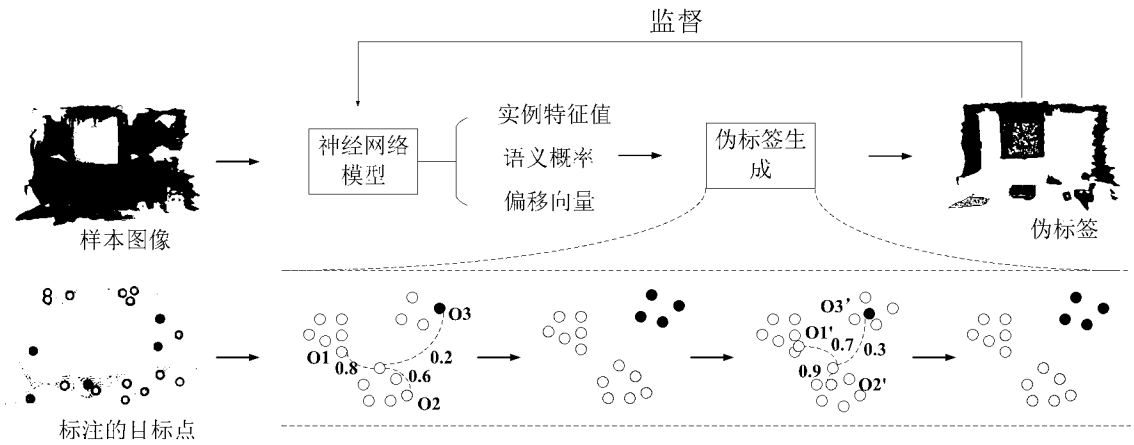

监督
样本图像
神经网络模型
实例特征值
语义概率
偏移向量
伪标签生成
伪标签
标注的目标点
O1
O2
O3
0.8
0.6
0.2
O1'
O2'
O3'
0.7
0.3
0.9

图 2

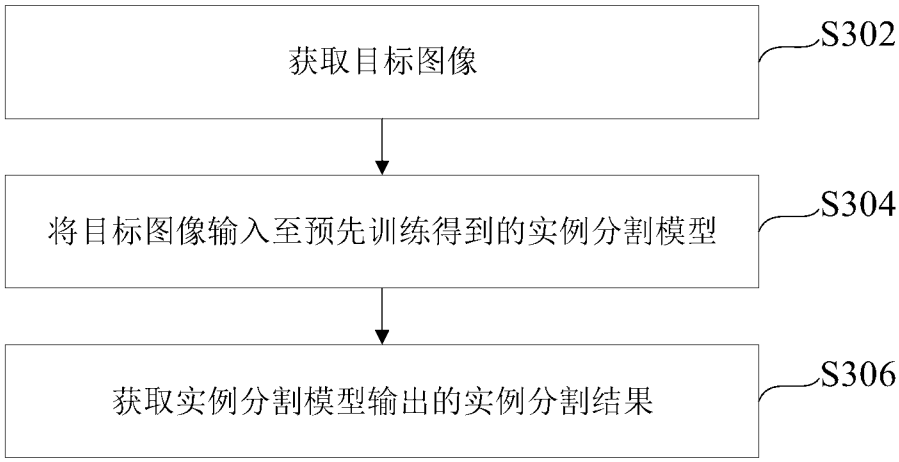

获取目标图像
S302
将目标图像输入至预先训练得到的实例分割模型
S304
获取实例分割模型输出的实例分割结果
S306

图 3

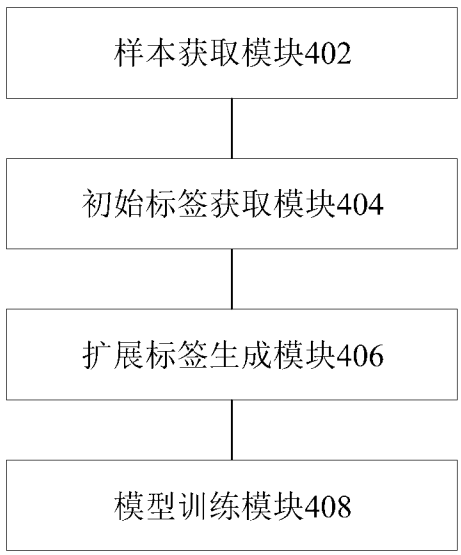

样本获取模块402
初始标签获取模块404
扩展标签生成模块406
模型训练模块408

图 4


图像获取模块502
模型输入模块504
结果获取模块506

图 5

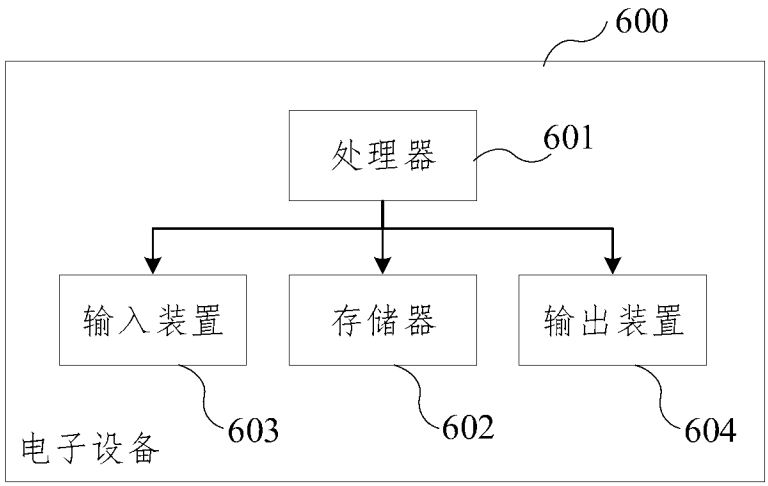


图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2023/107725

A. CLASSIFICATION OF SUBJECT MATTER

G06V10/82(2022.01)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)

IPC: G06V, 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)

CNTXT, CNABS, CNKI, ENTXT, ENTXTC, DWPI: 标签, 标注, 扩展, 实例分割, 增广, 样本, label, tag, expand, instance, segmentation, sample

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 115115923 A (BEIJING YOUZHUJU NETWORK TECHNOLOGY CO., LTD.) 27 September 2022 (2022-09-27) entire document	1-13
X	CN 112927244 A (TSINGHUA UNIVERSITY) 08 June 2021 (2021-06-08) description, paragraphs [0044]-[0058]	1, 9-13
A	US 2022207861 A1 (NEC CORP.) 30 June 2022 (2022-06-30) entire document	1-13

☐ 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
"D" document cited by the applicant in the international application
"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	Date of mailing of the international search report
09 October 2023	**19 October 2023**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2023/107725

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 115115923 A	27 September 2022	None	
CN 112927244 A	08 June 2021	None	
US 2022207861 A1	30 June 2022	JP 2022103136 A	07 July 2022

国际检索报告

国际申请号
PCT/CN2023/107725

A. 主题的分类

G06V10/82(2022.01)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)

IPC: G06V, 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））

CNTXT,CNABS,CNKI,ENTXT,ENTXTC,DWPI:标签,标注,扩展,实例分割,增广,样本,label, tag, expand, instance, segmentation, sample

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
PX	CN 115115923 A (北京有竹居网络技术有限公司) 2022年9月27日 (2022 - 09 - 27) 全文	1-13
X	CN 112927244 A (清华大学) 2021年6月8日 (2021 - 06 - 08) 说明书第44-58段	1，9-13
A	US 2022207861 A1 (NEC CORP) 2022年6月30日 (2022 - 06 - 30) 全文	1-13

☐ 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

“D” 申请人在国际申请中引证的文件

“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
2023年10月9日	2023年10月19日

ISA/CN的名称和邮寄地址	受权官员
中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088	白雪涛 电话号码 (+86) 62412069

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2023/107725

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 115115923 A	2022年9月27日	无	
CN 112927244 A	2021年6月8日	无	
US 2022207861 A1	2022年6月30日	JP 2022103136 A	2022年7月7日