

(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
2023 年 9 月 14 日 (14.09.2023)



(10) 国际公布号
WO 2023/169334 A1

(51) 国际专利分类号:
G06V 10/26 (2022.01)

(21) 国际申请号: PCT/CN2023/079602

(22) 国际申请日: 2023 年 3 月 3 日 (03.03.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202210225180.8 2022年3月9日 (09.03.2022) CN

(71) 申请人: 北京字跳网络技术有限公司
(**BEIJING ZITIAO NETWORK TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市海淀区紫金数码园 4 号楼 2 层 0207, Beijing 100190 (CN)。

(72) 发明人: 吴捷(**WU, Jie**); 中国北京市海淀区知春路 63 号中国卫星通信大厦今日头条小邮局, Beijing 100086 (CN)。 覃杰(**QIN, Jie**); 中国北京市海淀区知春路 63 号中国卫星通信大厦今日头条小邮局, Beijing 100086 (CN)。 肖学锋(**XIAO, Xuefeng**); 中国北京市海淀区知春路 63 号中国卫星通信大厦今日头条小邮局, Beijing 100086 (CN)。

(74) 代理人: 北京品源专利代理有限公司(**BEYOND ATTORNEYS AT LAW**); 中国北京市海淀区莲花池东路 39 号西金大厦 6 层, Beijing 100036 (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,

(54) **Title:** SEMANTIC SEGMENTATION METHOD AND APPARATUS FOR IMAGE, AND ELECTRONIC DEVICE AND STORAGE MEDIUM

(54) 发明名称: 图像的语义分割方法、装置、电子设备及存储介质

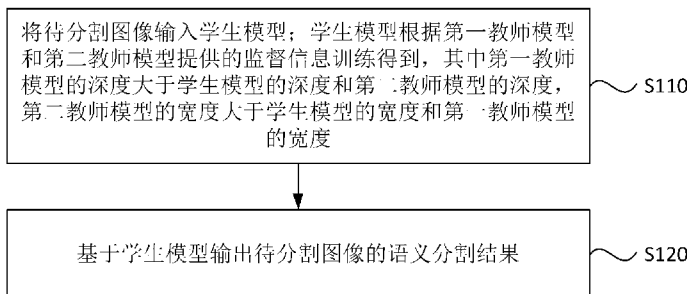


图 1

- S110 Input into a student model an image to be segmented, the student model being obtained by means of performing training according to supervision information provided by a first teacher model and a second teacher model, wherein the depth of the first teacher model is greater than the depth of the student model and the depth of the second teacher model, and the width of the second teacher model is greater than the width of the student model and the width of the first teacher model
- S120 Output a semantic segmentation result of said image on the basis of the student model

(57) **Abstract:** Provided in the present disclosure are a semantic segmentation method and apparatus for an image, and an electronic device and a storage medium. The method comprises: inputting into a student model an image to be segmented, the student model being obtained by means of performing training according to supervision information provided by a first teacher model and a second teacher model, wherein the depth of the first teacher model is greater than the depth of the student model and the depth of the second teacher model, and the width of the second teacher model is greater than the width of the student model and the width of the first teacher model;

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,
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, 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))。

and outputting a semantic segmentation result of said image on the basis of the student model.

(57) 摘要: 本公开提供了一种图像的语义分割方法、装置、电子设备及存储介质。该方法包括: 将待分割图像输入学生模型; 学生模型根据第一教师模型和第二教师模型提供的监督信息训练得到, 其中第一教师模型的深度大于学生模型的深度和第二教师模型的深度, 第二教师模型的宽度大于学生模型的宽度和第一教师模型的宽度; 基于学生模型输出待分割图像的语义分割结果。

图像的语义分割方法、装置、电子设备及存储介质

本申请要求在2022年03月09日提交中国专利局、申请号为202210225180.8的中国专利申请的优先权，该申请的全部内容通过引用结合在本申请中。

技术领域

本公开涉及计算机技术领域，例如涉及图像的语义分割方法、装置、电子设备及存储介质。

背景技术

图像语义分割技术是以语义属性作为划分标准实现逐像素分类预测的技术。

相关技术中，为保证语义分割效果，语义分割模型的深度、宽度通常较大。其中，模型的深度可以认为是模型的网络层数，模型的宽度可以认为是每层网络中的通道数。

相关技术的不足之处至少包括：应用大体量的语义分割模型需要以大量的资源作为代价，例如需要消耗大量计算资源和部署空间资源等。这对将语义分割模型部署至资源受限的设备带来巨大挑战。

发明内容

本公开提供了一种图像的语义分割方法、装置、电子设备及存储介质，能够在保证语义分割效果的基础上，采用轻量化模型实现图像语义分割，大大减少了资源耗费，有利于资源受限设备上的模型部署。

第一方面，本公开提供了一种图像的语义分割方法，包括：

将待分割图像输入学生模型；其中，所述学生模型根据第一教师模型和第二教师模型提供的监督信息训练得到，所述第一教师模型的深度大于所述学生模型的深度和所述第二教师模型的深度，所述第二教师模型的宽度大于所述学生模型的宽度和所述第一教师模型的宽度；

基于所述学生模型输出所述待分割图像的语义分割结果。

第二方面，本公开还提供了一种图像的语义分割装置，包括：

输入模块，设置为将待分割图像输入学生模型；其中，所述学生模型根据第一教师模型和第二教师模型提供的监督信息训练得到，所述第一教师模型的深度大于所述学生模型的深度和所述第二教师模型的深度，所述第二教师模型

的宽度大于所述学生模型的宽度和所述第一教师模型的宽度；

输出模块，设置为基于所述学生模型输出所述待分割图像的语义分割结果。

第三方面，本公开还提供了一种电子设备，所述电子设备包括：

一个或多个处理器；

存储装置，设置为存储一个或多个程序；

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现上述的图像的语义分割方法。

第四方面，本公开还提供了一种包含计算机可执行指令的存储介质，所述计算机可执行指令在由计算机处理器执行时用于执行上述的图像的语义分割方法。

第五方面，本公开还提供了一种计算机程序产品，包括承载在非暂态计算机可读介质上的计算机程序，所述计算机程序包含用于执行上述的图像的语义分割方法的程序代码。

附图说明

图1为本公开实施例一所提供的一种图像的语义分割方法的流程示意图；

图2为本公开实施例二所提供的一种图像的语义分割方法中学生模型的训练步骤的流程示意图；

图3为本公开实施例四所提供的一种图像的语义分割装置的结构示意图；

图4为本公开实施例五所提供的一种电子设备的结构示意图。

具体实施方式

下面将参照附图描述本公开的实施例。虽然附图中显示了一些实施例，然而本公开可以通过多种形式来实现，提供这些实施例是为了理解本公开。本公开的附图及实施例仅用于示例性作用。

本公开的方法实施方式中记载的多个步骤可以按照不同的顺序执行，和/或并行执行。此外，方法实施方式可以包括附加的步骤和/或省略执行示出的步骤。本公开的范围在此方面不受限制。

本文使用的术语“包括”及其变形是开放性包括，即“包括但不限于”。术语“基于”是“至少部分地基于”。术语“一个实施例”表示“至少一个实施例”；术语“另一实施例”表示“至少一个另外的实施例”；术语“一些实施例”表示“至少一些实施例”。其他术语的相关定义将在下文描述中给出。

本公开中提及的“第一”、“第二”等概念仅用于对不同的装置、模块或单元进行区分，并非用于限定这些装置、模块或单元所执行的功能的顺序或者相互依存关系。

本公开中提及的“一个”、“多个”的修饰是示意性而非限制性的，本领域技术人员应当理解，除非在上下文另有指出，否则应该理解为“一个或多个”。

实施例一

图 1 为本公开实施例一所提供的一种图像的语义分割方法流程示意图，本公开实施例适用于基于轻量化模型进行图像语义分割的情形。该方法可以由图像的语义分割装置来执行，该装置可以通过软件和/或硬件的形式实现，该装置可配置于电子设备中，例如配置于手机、电脑等电子设备中。

如图 1 所示，本实施例提供的图像的语义分割方法，可以包括：

S110、将待分割图像输入学生模型；学生模型根据第一教师模型和第二教师模型提供的监督信息训练得到，其中第一教师模型的深度大于学生模型的深度和第二教师模型的深度，第二教师模型的宽度大于学生模型的宽度和第一教师模型的宽度。

通过图像语义分割能够得到图像中每个对象的语义以及位置坐标，因此在围绕场景理解展开的诸多领域中具有巨大实用价值。针对不同领域，待分割图像各不相同。例如，在自动驾驶领域，待分割图像可以为实时的道路图像。通过对实时道路图像进行语义分割（例如对图像中的行人、车辆进行分割），可以为自动驾驶任务奠定坚实的基础。此外，本实施例的语义分割方法也可以对其他领域对应的待分割图像进行语义分割，在此不做穷举。

本公开实施例中，学生模型可以认为是较浅、较窄的轻量化模型；第一教师模型可以认为是较深、较窄的大体量模型；第二教师模型可以认为是较浅、较宽的大体量模型。其中，第一教师模型在深度维度上可以大于学生模型和第二教师模型，且第二教师模型的深度可以大于学生模型；第二教师模型在宽度维度上可以大于学生模型和第一教师模型，且第一教师模型的宽度可以不小于学生模型。其中，第一教师模型、第二教师模型和学生模型的深度、宽度的数值可以根据实际应用场景进行设置。例如，第一教师模型的深度可以为 101 层，第二教师模型的深度可以为 34 层，以及学生模型的深度可以为 17 层等。又如，第一教师模型的宽度可以等于学生模型，且可以为第二教师模型的宽度的一半。

第一教师模型和第二教师模型为两个互补的网络结构。更深的第一教师模型可以具有更好地提取高级语义和全局分类抽象的能力，这有助于在面向分类的任务中取得有效的结果。更宽的第二教师模型可以更擅于捕获多样化的局部

内容感知信息，这有利于对像素之间的上下文关系进行建模。基于这两个互补的教师模型监督学生模型训练，可以从更深、更宽的两个维度为学生模型提供全面的监督信息。更深维度的监督信息可增强学生模型的分类能力，更宽维度的监督信息可帮助学生模型建模像素间的上下文关系。通过充分利用大模型的性能优势完成知识蒸馏的过程，能够大大提升轻量化的学生模型的表现。

在一些实现方式中，第一教师模型和第二教师模型可以为预先训练的模型，且在学生模型训练过程中参数固定。

第一教师模型和第二教师模型可以预先经过全监督训练或半监督训练得到。由于全监督训练需要预先标注海量像素级的标签，可优先选用半监督的训练方式训练第一教师模型和第二教师模型。半监督训练方式，可以认为是利用少量有标签的图像和大量无标签的图像，对第一教师模型和第二教师模型进行训练。在半监督训练过程中，可以采用生成无标签数据的伪标签的方式，和/或采用一致性正则的方式，利用无标签数据缩小标签数据较少而带来的性能衰减。

在这些实现方式中，可以预先对第一教师模型和第二教师模型进行训练，并可以固定训练完成的第一教师模型的参数和第二教师模型的参数，以执行知识蒸馏过程来改进学生模型的性能。此外，在一些其他实现方式中，当利用有标签数据训练学生模型时，第一教师模型和第二教师模型也可以适当的进行参数调整，从而可在一定程度上使第一教师模型和第二教师模型在利用无标签数据训练学生模型时达到更好的监督效果。

S120、基于学生模型输出待分割图像的语义分割结果。

如果将传统大体量的语义分割模型中的主干网络直接用简化网络替代，将使语义分割性能急剧下降。相较于该方式，本公开实施例通过两个教师模型提供互为补充的监督信息来提高轻量化学生模型的性能，能够使轻量化的学生模型实现很好的语义分割性能的同时，保证很低的资源消耗量。由于学生模型具有很小的参数量和计算量，可以很方便地部署到资源受限的设备上。

本公开实施例提供的图像的语义分割方法，在数据集上进行了广泛实验，表明该方法具有有效性，为轻量级语义分割模型的训练开辟了先河。

在一些实现方式中，在将待分割图像输入学生模型之前，还包括：响应于本端设备的资源剩余量符合预设范围，将学生模型部署于本端设备中。

在实际的分割场景中，部署分割模型的电子设备的资源可能有限，例如手机端的计算资源可能有限。在本实现方式中，电子设备在部署语义分割模型之前，可以获取本端设备的资源剩余量，例如计算资源剩余量、存储资源剩余量等。若本端设备的资源剩余量符合预设范围，则可以认为本端设备当前可用的

资源有限。此时，可以获取并部署轻量化的学生模型至本端设备，从而可以实现资源受限设备上的模型部署。

此外，若本端设备资源剩余量超出预设范围，则可以认为本端设备当前可用的资源充裕。此时可部署的模型选择较为广泛，可以将本实施例提供的学生模型或传统的语义分割模型部署于本端设备中。

本公开实施例的技术方案，将待分割图像输入学生模型；学生模型根据第一教师模型和第二教师模型提供的监督信息训练得到，其中第一教师模型的深度大于学生模型的深度和第二教师模型的深度，第二教师模型的宽度大于学生模型的宽度和第一教师模型的宽度；基于学生模型输出待分割图像的语义分割结果。通过利用更深、更宽的两个教师模型，为轻量化的学生模型提供不同方面的监督信息，能够实现从两个复杂模型到简单模型的知识蒸馏，可以保证基于监督信息训练的学生模型具有较佳的语义分割效果。并且，轻量化的学生模型可以大大减少资源耗费，有利于资源受限设备上的模型部署。

实施例二

本公开实施例与上述实施例中所提供的图像的语义分割方法中多个方案可以结合。本实施例所提供的图像的语义分割方法，对基于监督信息训练学生模型的步骤进行了描述。通过根据学生模型和第一教师模型、第二教师模型的分割结果，可以确定学生模型的全局语义损失、局部特征损失和一致性损失。基于全局语义损失训练学生模型可帮助学生模型学习判别高层语义类别；基于局部特征损失训练学生模型可帮助学生模型捕获图像局部细节纹理的信息；基于一致性损失训练学生模型有利于实现同一输入的多个结果保持一致，从而提高语义分割精度。

在一些实现方式中，学生模型可以根据下述步骤训练得到：基于第一教师模型、第二教师模型和学生模型，分别输出样本图像的第一分割结果、第二分割结果和第三分割结果；根据第一分割结果、第二分割结果和第三分割结果确定学生模型的全局语义损失，局部特征损失和一致性损失；将所述全局语义损失、所述局部特征损失和所述一致性损失作为监督信息，训练所述学生模型。

模型的深度越大，可以具有更好地提取高级语义和全局分类抽象的能力。由于第一教师模型在深度维度上大于学生模型，可以根据第一教师模型的分割结果，为学生模型的分割结果提供更多的高级语义。并且，若第二教师模型的深度也大于学生模型，则也可以根据第二教师模型的分割结果，为学生模型的分割结果提供更多的高级语义。全局语义损失即可以认为是，第三分割结果与第一分割结果和/或第二分割结果相比，在高维语义特征方面的差异。

模型的宽度越大，可以更擅于捕获多样化的局部内容感知信息。由于第二教师模型在宽度维度上大于学生模型，可以根据第二教师模型在生成分割结果过程中的特征图的局部特征，为学生模型在生成分割结果过程中的特征图提供更细节的局部内容感知信息。并且，若第一教师模型的宽度也大于学生模型，则也可以根据第一教师模型在生成分割结果过程中的特征图的局部特征，为学生模型在生成分割结果过程中的特征图提供更细节的局部内容感知信息。局部特征损失即可以认为是，在生成第三分割结果过程中的特征图，与在生成第一分割结果和/或第二分割结果过程中的特征图相比，在局部上下文关系方面的差异。

在模型具备一定准确率的情况下，多模型针对同一图像的分割结果通常会趋于一致。由于第一教师模型和第二教师模型分别在全局分类抽象能力、捕获多样化的局部特征方面表现优于学生模型，可以根据第三分割结果与第一分割结果和第二分割结果的差异，确定一致性损失。

在这些实现方式中，基于全局语义损失训练学生模型可帮助学生模型学习判别高层语义类别；基于局部特征损失训练学生模型可帮助学生模型捕获图像局部细节纹理的信息；基于一致性损失训练学生模型有利于实现同一输入的多个结果保持一致，从而提高语义分割精度。

示例性的，图 2 为本公开实施例二所提供的一种图像的语义分割方法中学生模型的训练步骤的流程示意图。如图 2 所示，在一些实现方式中，学生模型可以根据下述步骤训练得到：

首先，可以基于第一教师模型 T_D 、第二教师模型 T_W 和学生模型 S ，分别输出样本图像的第一分割结果 Y^{T_D} 、第二分割结果 Y^{T_W} 和第三分割结果 Y^S 。

图 2 中可以看出，该训练过程采用的整体结构是一个三支的网络结构，由两个互补的大体量教师模型和一个轻量化的学生模型组成。其中，第一教师模型 T_D 的深度（图中用 Deep 表示）大于学生模型 S 的深度和第二教师模型 T_W 的深度，第二教师模型 T_W 的宽度（图中用 Wide 表示）大于学生模型 S 的宽度和第一教师模型 T_D 的宽度。

第一教师模型 T_D 可以为学生模型 S 提供全局的语义类别抽象概念，有利于学生模型 S 学习分类的能力；第二教师模型 T_W 利用更宽的通道数可以提取出更丰富的局部内容感知信息，辅助监督学生模型 S ，有助于学生模型 S 建模像素间的上下文关系。即，可以实现从两个复杂教师模型到简单学生模型的多粒度知识蒸馏，有利于突破轻量级模型的学习能力瓶颈，以保证基于监督信息训练的学生模型具有较佳的语义分割效果。

然后, 可以根据第三分割结果 Y^S 与第一分割结果 Y^{T_D} 的差异, 确定学生模型 S 的全局语义损失 (图中用 Global Semantic-sensitive Loss 表示)。该全局语义损失可以认为是由第一教师模型 T_D 提供给学生模型 S 的监督信息, 可用于表征更深的教师模型和学生模型之间的高维语义特征知识的差异。

也可以根据学生模型 S 确定的用于生成第三分割结果 Y^S 的特征图像 K_L^S , 与第二教师模型 T_W 确定的用于生成第二分割结果 Y^{T_W} 的特征图像 $K_L^{T_W}$ 之间的差异, 确定学生模型 S 的局部特征损失 (图中用 Local Content-aware Loss 表示)。该局部特征损失可以认为是由第二教师模型 T_W 提供给学生模型 S 的监督信息, 可用于表征更宽的教师模型和学生模型之间的局部上下文关系的差异。

还可以根据第三分割结果 Y^S 分别与第一分割结果 Y^{T_D} 和第二分割结果 Y^{T_W} 的差异, 确定学生模型 S 的一致性损失 (图中用 Complementary Consistency Loss 表示)。该一致性损失可以认为是第一教师模型 T_D 和第二教师模型 T_W 同时提供给学生模型 S 的监督信息。其中, 第一分割结果 Y^{T_D} 和第二分割结果 Y^{T_W} 中多个通道图像中的像素值可表征对应分割类别的概率值, 可以通过多个通道图像的像素值取最大值的方式, 分别得到第一伪标签 $Y_p^{T_D}$ 和第二伪标签 $Y_p^{T_W}$ 。通过第一伪标签 $Y_p^{T_D}$ 和第二伪标签 $Y_p^{T_W}$ 可辅助训练学生模型 S 。

上述全局语义损失、局部特征损失和一致性损失的求取步骤并无严格时序限制。例如, 可以在确定出第一分割结果 Y^{T_D} 、第二分割结果 Y^{T_W} 和第三分割结果 Y^S 之后, 同步计算该些损失; 又如, 可以在确定出特征图像 K_L^S 和 $K_L^{T_W}$ 时先计算局部特征损失, 在确定出第一分割结果 Y^{T_D} 、第二分割结果 Y^{T_W} 和第三分割结果 Y^S 之后, 再计算全局语义损失和一致性损失等。

最后, 将全局语义损失、局部特征损失和一致性损失作为监督信息, 训练学生模型 S 。

本公开实施例中, 在模型内解码器输出的特征层, 使用局部特征损失辅助监督学生网络, 在预测输出层使用全局语义损失提高学生网络的语义类别识别能力, 即实现了多层、多粒度的知识蒸馏方案来训练轻量化的学生模型, 以实现学生模型的高性能和低计算量。

再次参见图 2, 在一些实现方式中, 全局语义损失可以根据下述步骤确定: 将第一分割结果 Y^{T_D} 和第三分割结果 Y^S 进行逐通道池化 (例如可以为逐通道全局平均池化 (Global Average Pooling, GAP)), 分别得到第一全局向量 $K_G^{T_D}$ 和第二全局向量 K_G^S ; 将第一全局向量 $K_G^{T_D}$ 和第二全局向量 K_G^S 间多个维度的差值之和, 作为学生模型 S 的全局语义损失。

可以通过下述公式来确定第一全局向量：

$$K_G^{T_D} = G(Y^{T_D} \in R^{N \times H \times W});$$

其中， $Y^{T_D} \in R^{N \times H \times W}$ 表示第一分割结果 Y^{T_D} 的图像尺寸为 $N \times H \times W$ ； N 表示通道数量， H 表示图像高度， W 表示图像宽度，且后续相同格式上标所表示的含义与此处相同，将不再赘述。其中， $G(\cdot)$ 表示逐通道的全局平均池化运算，且运算得到的 $K_G^{T_D} \in R^{N \times 1 \times 1}$ ，可表示 N 个分割类别的全局语义类别向量。基于同种运算方式，可确定第二全局向量 K_G^S 。

可通过下述公式来确定全局语义损失：

$$L_{Sem}(K_G^S, K_G^{T_D}) = \sum_{i=1}^N (k_{Gi}^S - k_{Gi}^{T_D});$$

其中， $L_{Sem}(K_G^S, K_G^{T_D})$ 表示全局语义损失； k_{Gi}^S 和 $k_{Gi}^{T_D}$ 可以分别表示第二全局向量 K_G^S 和第一全局向量 $K_G^{T_D}$ 中第 i 个维度的数值； N 表示分割类别的总数量。

在这些实现方式中，通过全局语义损失可使学生模型尝试学习更高维的语义类别表示，这有助于在语义分割任务中为语义类别的判别提供全局指导。

再次参见图 2，在一些实现方式中，局部特征损失根据下述步骤确定：将第二教师模型 T_W 确定的特征图像 $K_L^{T_W}$ 和学生模型 S 确定的特征图像 K_L^S 进行逐通道、逐像素求取特征差值，并根据多个特征差值确定局部特征损失。

可以通过下述公式确定局部特征损失：

$$L_{Con}(K_L^S, K_L^{T_W}) = \frac{1}{C \times H \times W} \sum_{i=1}^C \sum_{j=1}^H \sum_{q=1}^W (k_{Lijq}^S - k_{Lijq}^{T_W})^2;$$

其中， $L_{Con}(K_L^S, K_L^{T_W})$ 表示局部特征损失； $C \times H \times W$ 表示特征图像 K_L^S 和特征图像 $K_L^{T_W}$ 的尺寸； k_{Lijq}^S 和 $k_{Lijq}^{T_W}$ 分别表示特征图像 K_L^S 和特征图像 $K_L^{T_W}$ 在第 i 个通道、第 j 个高度、第 q 个宽度上像素的特征值。

在这些实现方式中，局部特征损失旨在利用更宽的教师模型的通道优势来提供丰富的局部上下文信息，它可以提供辅助监督来指导学生模型对像素之间的上下文关系进行建模。

再次参见图 2，在一些实现方式中，一致性损失根据下述步骤确定：将第三分割结果 Y^S 分别与第一分割结果 Y^{T_D} 和第二分割结果 Y^{T_W} 的交叉熵损失之和，作为学生模型的一致性损失。

图中，第一伪标签 $Y_p^{T_D}$ 和第二伪标签 $Y_p^{T_W}$ 为基于第一分割结果 Y^{T_D} 和第二分割结果 Y^{T_W} 确定的伪标签。相应的，可以通过下述公式来确定一致性损失：

$$L_{Com}(Y, Y_p) = L_{ce}(Y, Y_p^{T_D}) + L_{ce}(Y, Y_p^{T_W})$$

$$= -\frac{1}{H \times W} \sum_{i=1}^{H \times W} [y_i \log(y_{pi}^{T_D}) + y_i \log(y_{pi}^{T_W})];$$

其中, Y^S 中多个通道图像中的像素值可表征对应分割类别的概率值, 可以通过多个通道图像的像素值取最大值的方式, 得到预测结果 Y 。其中, 一致性损失 $L_{Com}(Y, Y_p)$ 可以由 Y 和 $Y_p^{T_D}$ 的交叉熵损失 $L_{ce}(Y^S, Y_p^{T_D})$, 与 Y 和 $Y_p^{T_W}$ 的交叉熵损失 $L_{ce}(Y^S, Y_p^{T_W})$ 的和构成。其中, $H \times W$ 可以表示预测结果和两个伪标签的像素总数, y_i 、 $y_{pi}^{T_D}$ 和 $y_{pi}^{T_W}$ 可分别表示预测结果 Y 、第一伪标签 $Y_p^{T_D}$ 和第二伪标签 $Y_p^{T_W}$ 中第 i 个像素的预测的分割类别。此外, 除计算 Y 分别与 $Y_p^{T_D}$ 和 $Y_p^{T_W}$ 的交叉熵损失之外, 还可计算其他种类的图像间损失, 以确定一致性损失。

在这些实现方式中, 通过计算第一教师模型、第二教师模型和学生模型之间的互补的一致性损失, 可以保持同一输入的多个预测的一致性, 提高学生模型的性能。

本公开实施例的技术方案, 对基于监督信息训练学生模型的步骤进行了描述。通过根据学生模型和第一教师模型、第二教师模型的分割结果, 可以确定学生模型的全局语义损失、局部特征损失和一致性损失。基于全局语义损失训练学生模型可帮助学生模型学习判别高层语义类别; 基于局部特征损失训练学生模型可帮助学生模型捕获图像局部细节纹理的信息; 基于一致性损失训练学生模型有利于实现同一输入的多个结果保持一致, 从而提高语义分割精度。

此外, 本公开实施例提供的图像的语义分割方法与上述实施例提供的图像的语义分割方法属于同一构思, 未在本实施例中详尽描述的技术细节可参见上述实施例, 并且相同的技术特征在本实施例与上述实施例中具有相同的效果。

实施例三

本公开实施例与上述实施例中所提供的图像的语义分割方法中多个方案可以结合。本实施例所提供的图像的语义分割方法, 对样本图像为有标签样本图像时的监督信息进行了补充。通过根据学生模型的分割结果与标签之间的差异, 可以实现学生模型的有监督学习, 提高学生模型的语义分割精度。

在一些实现方式中, 若样本图像包括有标签的第一样本图像, 则学生模型的训练步骤, 还包括: 根据第三分割结果与第一样本图像的标签的差异, 确定学生模型的监督损失; 相应的, 将全局语义损失、局部特征损失和一致性损失作为监督信息, 训练学生模型, 包括: 将全局语义损失、局部特征损失、一致性损失和监督损失作为监督信息, 训练学生模型。

当样本图像仅包含有标签的第一样本图像时, 学生模型的训练方式可认为是全监督训练; 当样本图像既包含有标签的第一样本图像, 又包含无标签的第

二样本图像时，学生模型的训练方式可认为是半监督训练。以半监督训练方式训练学生模型时，可以根据第一教师模型和第二教师模型输出的预测结果确定伪标签，对学生模型进行训练。

当样本图像包含第一样本图像时，除了可以确定全局语义损失、局部特征损失和一致性损失之外，还可以确定监督损失。并且，可以结合上述损失训练学生模型，以提高学生模型精度。

监督损失根据下述步骤可以确定：将第三分割结果与第一样本图像的标签的交叉熵损失，作为学生模型的监督损失。可以通过下述公式来确定监督损失：

$$L_{Sup}^l(Y, \hat{Y}) = -\frac{1}{H \times W} \sum_{i=1}^{H \times W} [y_i \log(\hat{y}_i)];$$

其中， $L_{Sup}^l(Y, \hat{Y})$ 可以表示学生模型的预测结果 Y 与第一样本图像的标签 $\hat{Y}$ 间的监督损失，且上标 l 可表示该监督损失在有标签的样本数据时确定。其中， $H \times W$ 可以表示预测结果 Y 和标签 $\hat{Y}$ 的像素总数， y_i 和 $\hat{y}_i$ 可分别表示预测结果 Y 和标签 $\hat{Y}$ 中第 i 个像素的预测的分割类别。此外，除计算 Y 与 $\hat{Y}$ 的交叉熵损失之外，还可计算其他种类的图像间损失，以确定监督损失。

此外，针对第一样本图像和第二样本图像，可以从上述损失中分别选用不同的损失训练学生模型。

示例性的，采用第一样本图像训练学生模型时，可以选用监督损失和一致性损失训练学生模型；采用第二样本图像训练学生模型时，可以选用全局语义损失、局部特征损失和一致性损失训练学生模型。

此时，学生模型训练的总损失可以用公式 $L = L_{Sup}^l + L_{Com}^{u,l} + \lambda_1 L_{Sem}^u + \lambda_2 L_{Con}^u$ 表示；其中， L 可以为总损失； L_{Sup}^l 可以表示有标签的第一样本图像对应的监督损失； $L_{Com}^{u,l}$ 可以表示有标签的第一样本图像对应的一致性损失 L_{Com}^l 和无标签的第二样本图像对应的一致性损失 L_{Com}^u 的综合损失； L_{Sem}^u 可以表示无标签的第二样本图像对应的全局语义损失； L_{Con}^u 可以表示无标签的第二样本图像对应的局部特征损失； λ_1 和 λ_2 是损失函数的加权参数，且该两参数可以根据经验值或实验值进行设置。

本公开实施例的技术方案，对样本图像为有标签样本图像时的监督信息进行了补充。通过根据学生模型的分割结果与标签之间的差异，可以实现学生模型的有监督学习，提高学生模型的语义分割精度。此外，本公开实施例提供的图像的语义分割方法与上述实施例提供的图像的语义分割方法属于同一构思，未在本实施例中详尽描述的技术细节可参见上述实施例，并且相同的技术特征在本实施例与上述实施例中具有相同的效果。

实施例四

图 3 为本公开实施例四所提供的一种图像的语义分割装置的结构示意图。本公开实施例适用于基于轻量化模型进行图像语义分割的情形。

如图 3 所示，本实施例提供的图像的语义分割装置，可以包括：

输入模块 310，设置为将待分割图像输入学生模型；学生模型根据第一教师模型和第二教师模型提供的监督信息训练得到，其中第一教师模型的深度大于学生模型的深度和第二教师模型的深度，第二教师模型的宽度大于学生模型的宽度和第一教师模型的宽度；输出模块 320，设置为基于学生模型输出待分割图像的语义分割结果。

在一些实现方式中，图像的语义分割装置，可以包括：

模型训练模块，可以设置为根据下述步骤训练得到学生模型：

基于第一教师模型、第二教师模型和学生模型，分别输出样本图像的第一分割结果、第二分割结果和第三分割结果；根据第一分割结果、第二分割结果和第三分割结果确定学生模型的全局语义损失，局部特征损失和一致性损失；将全局语义损失、局部特征损失和一致性损失作为监督信息，训练学生模型。

在一些实现方式中，模型训练模块，可以设置为：

根据第三分割结果与第一分割结果的差异，确定学生模型的全局语义损失；根据学生模型确定的用于生成第三分割结果的特征图像，与第二教师模型确定的用于生成第二分割结果的特征图像之间的差异，确定学生模型的局部特征损失；根据第三分割结果分别与第一分割结果和第二分割结果的差异，确定学生模型的一致性损失。

在一些实现方式中，模型训练模块，可以设置为根据下述步骤确定全局语义损失：

将第一分割结果和第三分割结果进行逐通道池化，分别得到第一全局向量和第二全局向量；将第一全局向量和第二全局向量间多个维度的差值之和，作为学生模型的全局语义损失。

在一些实现方式中，模型训练模块，可以设置为根据下述步骤确定局部特征损失：

将第二教师模型确定的特征图像和学生模型确定的特征图像进行逐通道、逐像素求取特征差值，并根据多个特征差值确定局部特征损失。

在一些实现方式中，模型训练模块，可以设置为根据下述步骤确定一致性损失：

将第三分割结果分别与第一分割结果和第二分割结果的交叉熵损失之和，

作为学生模型的一致性损失。

在一些实现方式中，若样本图像包括有标签的第一样本图像，则模型训练模块，还可以设置为：

根据第三分割结果与第一样本图像的标签的差异，确定学生模型的监督损失；相应的，模型训练模块，可以设置为：

将全局语义损失、局部特征损失、一致性损失和监督损失作为监督信息，训练学生模型。

在一些实现方式中，模型训练模块，可以设置为根据下述步骤确定监督损失：

将第三分割结果与第一样本图像的标签的交叉熵损失，作为学生模型的监督损失。

在一些实现方式中，第一教师模型和第二教师模型为预先训练的模型，且在学生模型训练过程中参数固定。

在一些实现方式中，图像的语义分割装置，还可以包括：

部署模块，可以设置为在将待分割图像输入学生模型之前，响应于本端设备的资源剩余量符合预设范围，将学生模型部署于本端设备中。

本公开实施例所提供的图像的语义分割装置，可执行本公开任意实施例所提供的图像的语义分割方法，具备执行方法相应的功能模块和有益效果。

上述装置所包括的多个单元和模块只是按照功能逻辑进行划分的，但并不局限于上述的划分，只要能够实现相应的功能即可；另外，多个功能单元的名称也只是为了便于相互区分，并不用于限制本公开实施例的保护范围。

实施例五

下面参考图 4，其示出了适于用来实现本公开实施例的电子设备（例如图 4 中的终端设备或服务器）400 的结构示意图。本公开实施例中的终端设备可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、个人数字助理（Personal Digital Assistant, PDA）、平板电脑（Portable Android Device, PAD）、便携式多媒体播放器（Portable Media Player, PMP）、车载终端（例如车载导航终端）等等的移动终端以及诸如数字电视（Television, TV）、台式计算机等等的固定终端。图 4 示出的电子设备 400 仅仅是一个示例，不应对本公开实施例的功能和使用范围带来任何限制。

如图 4 所示，电子设备 400 可以包括处理装置（例如中央处理器、图形处理器等）401，其可以根据存储在只读存储器（Read-Only Memory, ROM）402

中的程序或者从存储装置 408 加载到随机访问存储器(Random Access Memory, RAM) 403 中的程序而执行多种适当的动作和处理。在 RAM 403 中, 还存储有电子设备 400 操作所需的多种程序和数据。处理装置 401、ROM 402 以及 RAM 403 通过总线 404 彼此相连。输入/输出 (Input/Output, I/O) 接口 405 也连接至总线 404。

通常, 以下装置可以连接至 I/O 接口 405: 包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置 406; 包括例如液晶显示器 (Liquid Crystal Display, LCD)、扬声器、振动器等的输出装置 407; 包括例如磁带、硬盘等的存储装置 408; 以及通信装置 409。通信装置 409 可以允许电子设备 400 与其他设备进行无线或有线通信以交换数据。虽然图 4 示出了具有多种装置的电子设备 400, 并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。

根据本公开的实施例, 上文参考流程图描述的过程可以被实现为计算机软件程序。例如, 本公开的实施例包括一种计算机程序产品, 其包括承载在非暂态计算机可读介质上的计算机程序, 该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中, 该计算机程序可以通过通信装置 409 从网络上被下载和安装, 或者从存储装置 408 被安装, 或者从 ROM 402 被安装。在该计算机程序被处理装置 401 执行时, 执行本公开实施例的图像的语义分割方法中限定的上述功能。

本公开实施例提供的电子设备与上述实施例提供的图像的语义分割方法属于同一构思, 未在本实施例中详尽描述的技术细节可参见上述实施例, 并且本实施例与上述实施例具有相同的效果。

实施例六

本公开实施例提供了一种计算机存储介质, 其上存储有计算机程序, 该程序被处理器执行时实现上述实施例所提供的图像的语义分割方法。

本公开上述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——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。计算机可读存储介质的例子可以包括但不限于: 具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、RAM、ROM、可擦式可编程只读存储器(Erasable Programmable Read-Only Memory, EPROM)或闪存(FLASH)、光纤、便携式紧凑磁盘只读存储器(Compact Disc Read-Only Memory, CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中, 计算机可读存储介质可以是任何包含或存储程序的有形介质, 该程序可以被指令执行

系统、装置或者器件使用或者与其结合使用。而在本公开中，计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：电线、光缆、射频（Radio Frequency, RF）等等，或者上述的任意合适的组合。

在一些实施方式中，客户端、服务器可以利用诸如超文本传输协议（Hyper Text Transfer Protocol, HTTP）之类的任何当前已知或未来研发的网络协议进行通信，并且可以与任意形式或介质的数字数据通信（例如，通信网络）互连。通信网络的示例包括局域网（Local Area Network, LAN），广域网（Wide Area Network, WAN），网际网（例如，互联网）以及端对端网络（例如，ad hoc 端对端网络），以及任何当前已知或未来研发的网络。

上述计算机可读介质可以是上述电子设备中所包含的；也可以是单独存在，而未装配入该电子设备中。

上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被该电子设备执行时，使得该电子设备：

将待分割图像输入学生模型；学生模型根据第一教师模型和第二教师模型提供的监督信息训练得到，其中第一教师模型的深度大于学生模型的深度和第二教师模型的深度，第二教师模型的宽度大于学生模型的宽度和第一教师模型的宽度；基于学生模型输出待分割图像的语义分割结果。

可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的计算机程序代码，上述程序设计语言包括但不限于面向对象的程序设计语言——诸如 Java、Smalltalk、C++，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机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括 LAN 或 WAN——连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。

附图中的流程图和框图，图示了按照本公开多种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图

中的每个方框可以代表一个模块、程序段、或代码的一部分，该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的是，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

描述于本公开实施例中涉及到的单元可以通过软件的方式实现，也可以通过硬件的方式来实现。其中，单元、模块的名称在一种情况下并不构成对该单元、模块本身的限定。

本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例如，非限制性地，可以使用的示范类型的硬件逻辑部件包括：现场可编程门阵列(Field Programmable Gate Array, FPGA)、专用集成电路(Application Specific Integrated Circuit, ASIC)、专用标准产品(Application Specific Standard Parts, ASSP)、片上系统(System on Chip, SOC)、复杂可编程逻辑设备(Complex Programming Logic Device, CPLD)等等。

在本公开的上下文中，机器可读介质可以是有形的介质，其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备，或者上述内容的任何合适组合。机器可读存储介质的示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、RAM、ROM、EPROM或快闪存储器、光纤、CD-ROM、光学储存设备、磁储存设备、或上述内容的任何合适组合。

根据本公开的一个或多个实施例，【示例一】提供了一种图像的语义分割方法，该方法包括：

将待分割图像输入学生模型；所述学生模型根据第一教师模型和第二教师模型提供的监督信息训练得到，其中所述第一教师模型的深度大于所述学生模型的深度和所述第二教师模型的深度，所述第二教师模型的宽度大于所述学生模型的宽度和所述第一教师模型的宽度；

基于所述学生模型输出所述待分割图像的语义分割结果。

根据本公开的一个或多个实施例，【示例二】提供了一种图像的语义分割

方法，还包括：

在一些实现方式中，所述学生模型根据下述步骤训练得到：

基于所述第一教师模型、所述第二教师模型和所述学生模型，分别输出样本图像的第一分割结果、第二分割结果和第三分割结果；

根据所述第一分割结果、第二分割结果和第三分割结果确定学生模型的全局语义损失，局部特征损失和一致性损失；

将所述全局语义损失、所述局部特征损失和所述一致性损失作为监督信息，训练所述学生模型。

根据本公开的一个或多个实施例，【示例三】提供了一种图像的语义分割方法，还包括：

在一些实现方式中，所述根据所述第一分割结果、第二分割结果和第三分割结果确定学生模型的全局语义损失，局部特征损失和一致性损失，包括：

根据所述第三分割结果与所述第一分割结果的差异，确定所述学生模型的全局语义损失；

根据所述学生模型确定的用于生成所述第三分割结果的特征图像，与所述第二教师模型确定的用于生成所述第二分割结果的特征图像之间的差异，确定所述学生模型的局部特征损失；

根据所述第三分割结果分别与所述第一分割结果和所述第二分割结果的差异，确定所述学生模型的一致性损失。

根据本公开的一个或多个实施例，【示例四】提供了一种图像的语义分割方法，还包括：

在一些实现方式中，所述全局语义损失根据下述步骤确定：

将所述第一分割结果和所述第三分割结果进行逐通道池化，分别得到第一全局向量和第二全局向量；

将所述第一全局向量和所述第二全局向量间多个维度的差值之和，作为所述学生模型的全局语义损失。

根据本公开的一个或多个实施例，【示例五】提供了一种图像的语义分割方法，还包括：

在一些实现方式中，所述局部特征损失根据下述步骤确定：

将所述第二教师模型确定的特征图像和所述学生模型确定的特征图像进行逐通道、逐像素求取特征差值，并根据多个特征差值确定局部特征损失。

根据本公开的一个或多个实施例，【示例六】提供了一种图像的语义分割方法，还包括：

在一些实现方式中，所述一致性损失根据下述步骤确定：

将所述第三分割结果分别与所述第一分割结果和所述第二分割结果的交叉熵损失之和，作为所述学生模型的一致性损失。

根据本公开的一个或多个实施例，【示例七】提供了一种图像的语义分割方法，还包括：

在一些实现方式中，若所述样本图像包括有标签的第一样本图像，则所述学生模型的训练步骤，还包括：

根据所述第三分割结果与所述第一样本图像的标签的差异，确定所述学生模型的监督损失；

相应的，所述将所述全局语义损失、所述局部特征损失和一致性损失作为监督信息，训练所述学生模型，包括：

将所述全局语义损失、所述局部特征损失、所述一致性损失和所述监督损失作为监督信息，训练所述学生模型。

根据本公开的一个或多个实施例，【示例八】提供了一种图像的语义分割方法，还包括：

在一些实现方式中，所述监督损失根据下述步骤确定：

将所述第三分割结果与所述第一样本图像的标签的交叉熵损失，作为所述学生模型的监督损失。

根据本公开的一个或多个实施例，【示例九】提供了一种图像的语义分割方法，还包括：

在一些实现方式中，所述第一教师模型和所述第二教师模型为预先训练的模型，且在所述学生模型训练过程中参数固定。

根据本公开的一个或多个实施例，【示例十】提供了一种图像的语义分割方法，还包括：

在一些实现方式中，在所述将待分割图像输入学生模型之前，还包括：

响应于本端设备的资源剩余量符合预设范围，将所述学生模型部署于所述本端设备中。

此外，虽然采用特定次序描绘了多个操作，但是这不应当理解为要求这些操作以所示出的特定次序或以顺序次序执行来执行。在一定环境下，多任务和

并行处理可能是有利的。同样地，虽然在上面论述中包含了多个实现细节，但是这些不应当被解释为对本公开的范围的限制。在单独的实施例的上下文中描述的一些特征还可以组合地实现在单个实施例中。相反地，在单个实施例的上下文中描述的多种特征也可以单独地或以任何合适的子组合的方式实现在多个实施例中。

权 利 要 求 书

1. 一种图像的语义分割方法，包括：

将待分割图像输入学生模型；其中，所述学生模型根据第一教师模型和第二教师模型提供的监督信息训练得到，所述第一教师模型的深度大于所述学生模型的深度和所述第二教师模型的深度，所述第二教师模型的宽度大于所述学生模型的宽度和所述第一教师模型的宽度；

基于所述学生模型输出所述待分割图像的语义分割结果。

2. 根据权利要求 1 所述的方法，其中，所述学生模型根据下述方法训练得到：

基于所述第一教师模型、所述第二教师模型和所述学生模型，分别输出样本图像的第一分割结果、第二分割结果和第三分割结果；

根据所述第一分割结果、所述第二分割结果和所述第三分割结果确定所述学生模型的全局语义损失，局部特征损失和一致性损失；

将所述全局语义损失、所述局部特征损失和所述一致性损失作为监督信息，训练所述学生模型。

3. 根据权利要求 2 所述的方法，其中，所述根据所述第一分割结果、所述第二分割结果和所述第三分割结果确定所述学生模型的全局语义损失，局部特征损失和一致性损失，包括：

根据所述第三分割结果与所述第一分割结果的差异，确定所述学生模型的全局语义损失；

根据所述学生模型确定的用于生成所述第三分割结果的特征图像，与所述第二教师模型确定的用于生成所述第二分割结果的特征图像之间的差异，确定所述学生模型的局部特征损失；

根据所述第三分割结果分别与所述第一分割结果和所述第二分割结果的差异，确定所述学生模型的一致性损失。

4. 根据权利要求 3 所述的方法，其中，所述根据所述第三分割结果与所述第一分割结果的差异，确定所述学生模型的全局语义损失，包括：

将所述第一分割结果和所述第三分割结果进行逐通道池化，分别得到第一全局向量和第二全局向量；

将所述第一全局向量和所述第二全局向量间多个维度的差值之和，作为所述学生模型的全局语义损失。

5. 根据权利要求 3 所述的方法，其中，所述根据所述学生模型确定的用于

生成所述第三分割结果的特征图像，与所述第二教师模型确定的用于生成所述第二分割结果的特征图像之间的差异，确定所述学生模型的局部特征损失，包括：

将所述第二教师模型确定的特征图像和所述学生模型确定的特征图像进行逐通道、逐像素求取特征差值，并根据多个特征差值确定局部特征损失。

6. 根据权利要求 3 所述的方法，其中，所述根据所述第三分割结果分别与所述第一分割结果和所述第二分割结果的差异，确定所述学生模型的一致性损失，包括：

将所述第三分割结果分别与所述第一分割结果和所述第二分割结果的交叉熵损失之和，作为所述学生模型的一致性损失。

7. 根据权利要求 2 所述的方法，其中，在所述样本图像包括有标签的第一样本图像的情况下，所述学生模型的训练方法，还包括：

根据所述第三分割结果与所述第一样本图像的标签的差异，确定所述学生模型的监督损失；

所述将所述全局语义损失、所述局部特征损失和所述一致性损失作为监督信息，训练所述学生模型，包括：

将所述全局语义损失、所述局部特征损失、所述一致性损失和所述监督损失作为监督信息，训练所述学生模型。

8. 根据权利要求 7 所述的方法，其中，所述根据所述第三分割结果与所述第一样本图像的标签的差异，确定所述学生模型的监督损失，包括：

将所述第三分割结果与所述第一样本图像的标签的交叉熵损失，作为所述学生模型的监督损失。

9. 根据权利要求 1-8 中任一所述的方法，其中，所述第一教师模型和所述第二教师模型为预先训练的模型，且在所述学生模型训练过程中参数固定。

10. 根据权利要求 1-8 中任一所述的方法，在所述将待分割图像输入学生模型之前，还包括：

响应于本端设备的资源剩余量符合预设范围，将所述学生模型部署于所述本端设备中。

11. 一种图像的语义分割装置，包括：

输入模块，设置为将待分割图像输入学生模型；其中，所述学生模型根据第一教师模型和第二教师模型提供的监督信息训练得到，所述第一教师模型的

深度大于所述学生模型的深度和所述第二教师模型的深度，所述第二教师模型的宽度大于所述学生模型的宽度和所述第一教师模型的宽度；

输出模块，设置为基于所述学生模型输出所述待分割图像的语义分割结果。

12. 一种电子设备，包括：

至少一个处理器；

存储装置，设置为存储至少一个程序；

当所述至少一个程序被所述至少一个处理器执行，使得所述至少一个处理器实现如权利要求 1-10 中任一所述的图像的语义分割方法。

13. 一种包含计算机可执行指令的存储介质，所述计算机可执行指令在由计算机处理器执行时用于执行如权利要求 1-10 中任一所述的图像的语义分割方法。

14. 一种计算机程序产品，包括承载在非暂态计算机可读介质上的计算机程序，所述计算机程序包含用于执行如权利要求 1-10 中任一所述的图像的语义分割方法的程序代码。

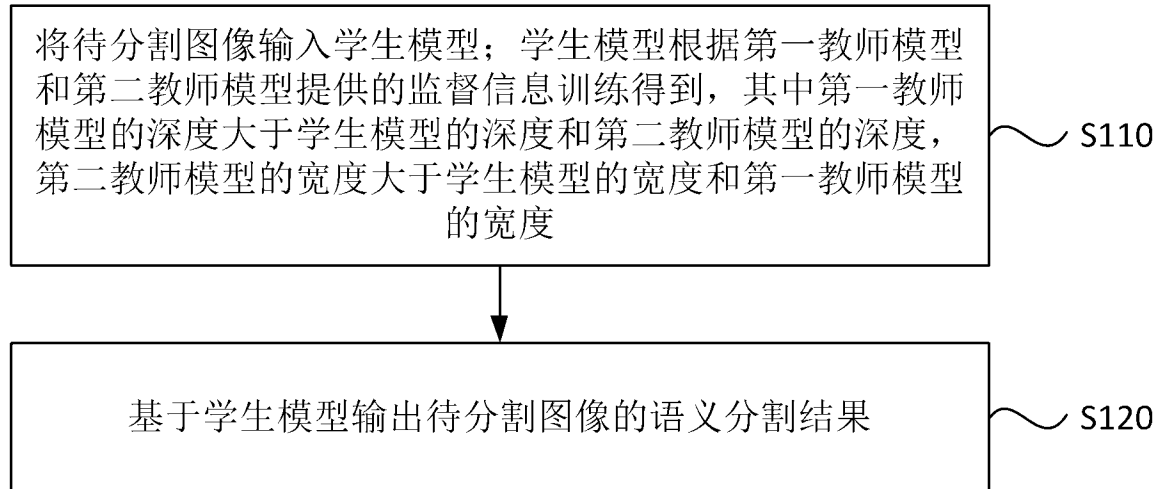


图 1

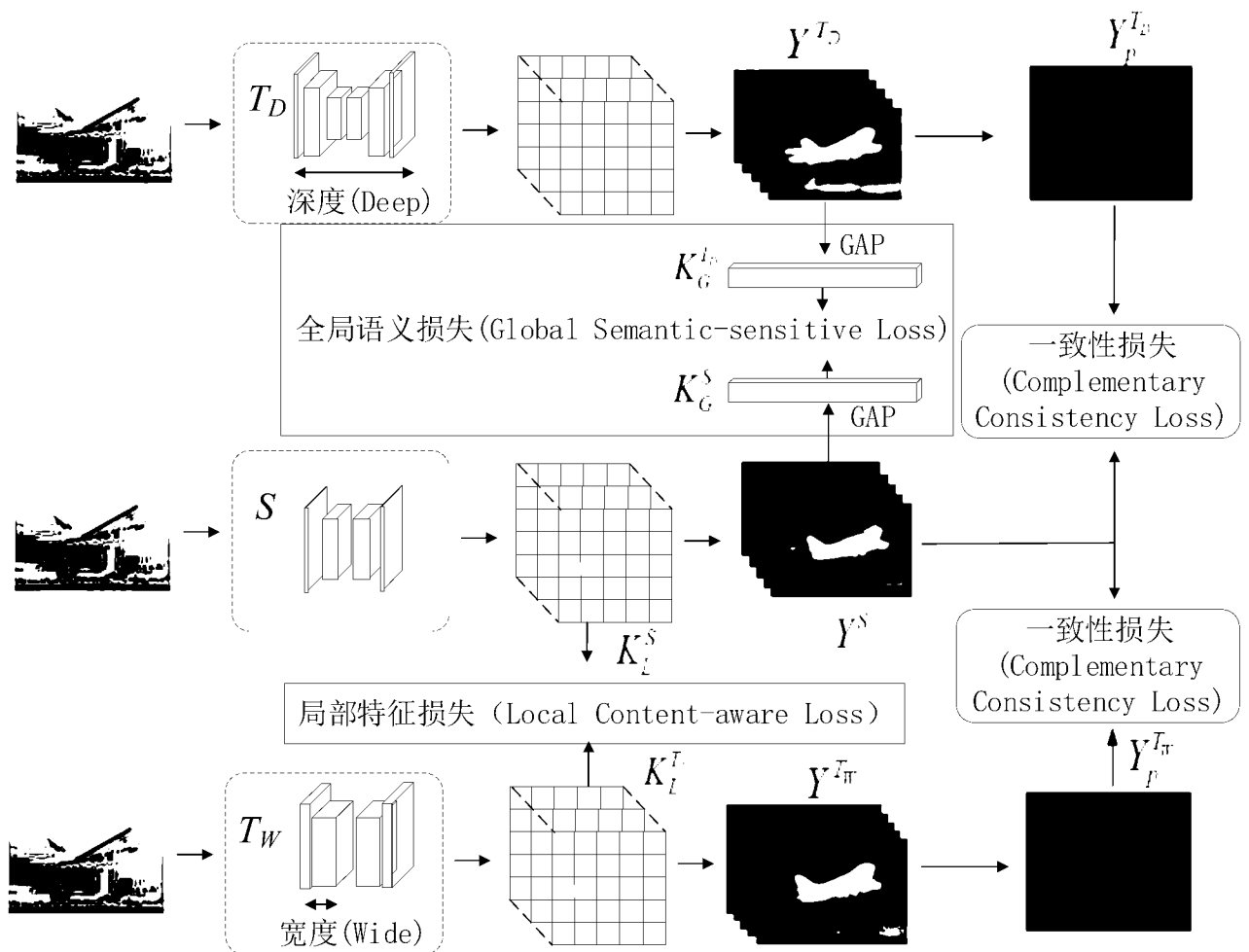


图 2

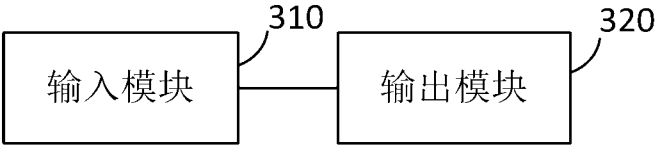


图 3

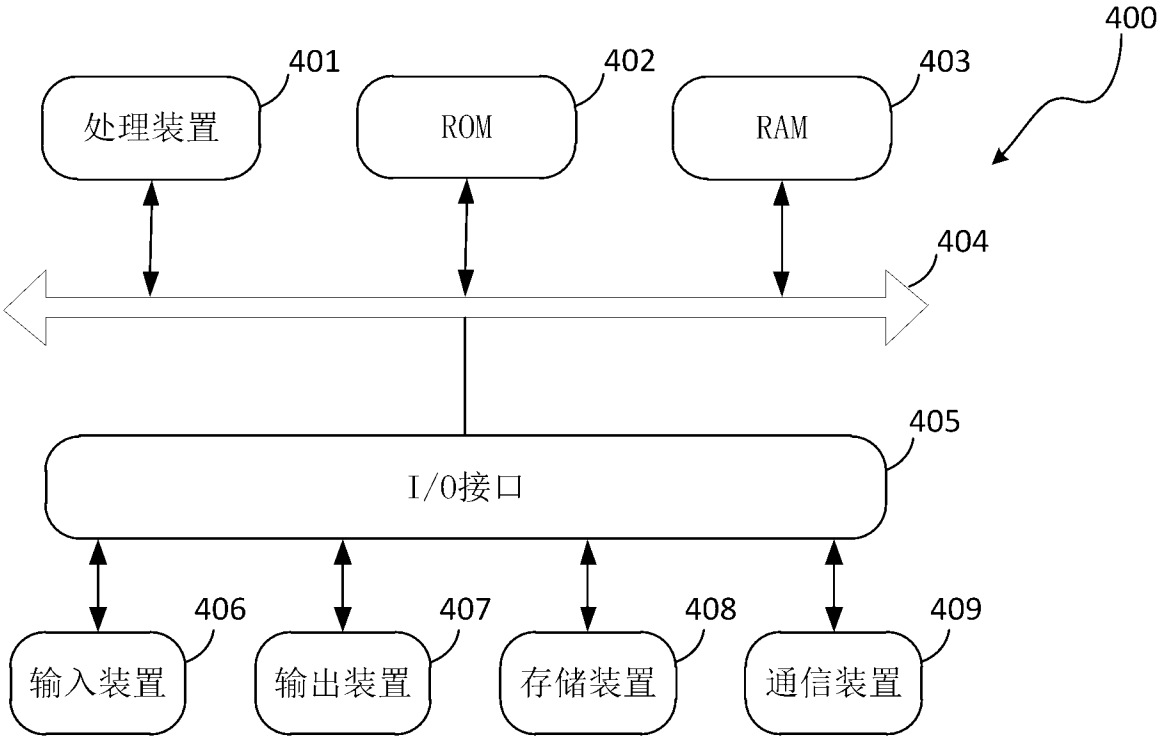


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/079602

A. CLASSIFICATION OF SUBJECT MATTER

G06V10/26(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: 层数, 教师, 学生, 模型, 宽度, 厚度, 深度, 通道数, 语义分割, teacher, student, model, width, depth, number, channel, semantic, segment+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 113449851 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 28 September 2021 (2021-09-28) description, paragraphs [0111]-[0118], and figure 5	1, 9-14
Y	CN 113627545 A (SHANDONG UNIVERSITY) 09 November 2021 (2021-11-09) description, paragraphs [0040]-[0049]	1, 9-14
Y	CN 113705362 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 26 November 2021 (2021-11-26) description, paragraphs [0035]-[0044]	1, 9-14
A	CN 111950638 A (XIAMEN MEITUZHILIA TECHNOLOGY CO., LTD.) 17 November 2020 (2020-11-17) entire document	1-14
A	KR 102225579 B1 (UNIV AJOU IND ACADEMIC COOP FOUND) 10 March 2021 (2021-03-10) entire document	1-14



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

13 June 2023

Date of mailing of the international search report

19 June 2023

Name and mailing address of the ISA/CN

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

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/079602

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	113449851	A	28 September 2021	WO	2023284416	A1	19 January 2023
CN	113627545	A	09 November 2021	None			
CN	113705362	A	26 November 2021	None			
CN	111950638	A	17 November 2020	None			
KR	102225579	B1	10 March 2021	None			

国际检索报告

国际申请号

PCT/CN2023/079602

A. 主题的分类

G06V10/26 (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: 层数, 教师, 学生, 模型, 宽度, 厚度, 深度, 通道数, 语义分割, teacher, student, model, width, depth, number, channel, semantic, segment+

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
X	CN 113449851 A (北京字跳网络技术有限公司) 2021年9月28日 (2021 - 09 - 28) 说明书第[0111]-[0118]段, 附图5	1, 9-14
Y	CN 113627545 A (山东大学) 2021年11月9日 (2021 - 11 - 09) 说明书第[0040]-[0049]段	1, 9-14
Y	CN 113705362 A (北京百度网讯科技有限公司) 2021年11月26日 (2021 - 11 - 26) 说明书第[0035]-[0044]段	1, 9-14
A	CN 111950638 A (厦门美图之家科技有限公司) 2020年11月17日 (2020 - 11 - 17) 全文	1-14
A	KR 102225579 B1 (UNIV AJOU IND ACADEMIC COOP FOUND) 2021年3月10日 (2021 - 03 - 10) 全文	1-14

☐ 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

2023年6月13日

国际检索报告邮寄日期

2023年6月19日

ISA/CN的名称和邮寄地址

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

授权官员

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/CN2023/079602

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	113449851	A	2021年9月28日	W0	2023284416	A1	2023年1月19日
CN	113627545	A	2021年11月9日	无			
CN	113705362	A	2021年11月26日	无			
CN	111950638	A	2020年11月17日	无			
KR	102225579	B1	2021年3月10日	无			