

(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
2025 年 1 月 2 日 (02.01.2025)



(10) 国际公布号
WO 2025/001002 A1

- (51) 国际专利分类号:
G06V 10/26 (2022.01) **G06V 10/774** (2022.01)
G06V 20/70 (2022.01) **G06V 10/80** (2022.01)
- (21) 国际申请号: PCT/CN2023/141742
- (22) 国际申请日: 2023 年 12 月 26 日 (26.12.2023)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202310781685.7 2023年6月28日 (28.06.2023) CN
- (71) 申请人: 深圳市优必选科技股份有限公司(**UBTECH ROBOTICS CORP LTD**) [CN/CN]; 中国广东省深圳市南山区桃源街道长源社区学苑大道1001号南山智园C1栋2201, Guangdong 518000 (CN)。

- (72) 发明人: 潘柏宇 (**PAN, Baiyu**); 中国广东省深圳市南山区桃源街道长源社区学苑大道1001号南山智园C1栋2201, Guangdong 518000 (CN)。 焦继超 (**JIAO, Jichao**); 中国广东省深圳市南山区桃源街道长源社区学苑大道1001号南山智园C1栋2201, Guangdong 518000 (CN)。
- (74) 代理人: 深圳中一联合知识产权代理有限公司 (**SHENZHEN ZHONGYI UNION INTELLECTUAL PROPERTY AGENCY CO., LTD.**); 中国广东省深圳市福田区莲花街道紫荆社区深南大道6008号深圳特区报业大厦33层, Guangdong 518034 (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,

(54) **Title:** SEMANTIC SEGMENTATION METHOD AND APPARATUS, AND COMPUTER-READABLE STORAGE MEDIUM AND ROBOT

(54) 发明名称: 一种语义分割方法、装置、计算机可读存储介质及机器人

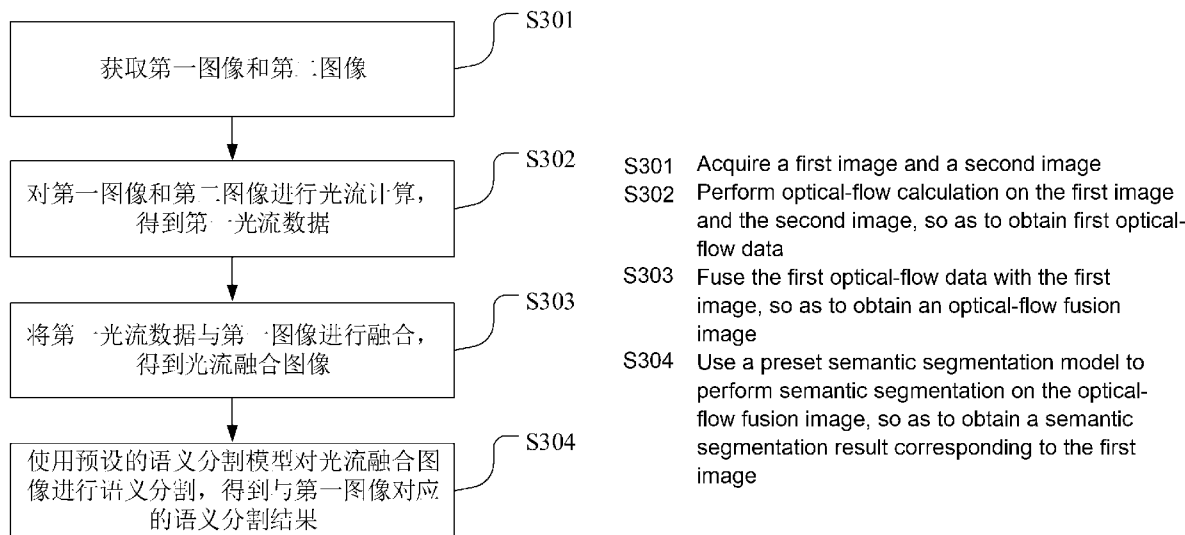


图 3

(57) **Abstract:** The present application belongs to the technical field of robots, and in particular relates to a semantic segmentation method and apparatus, and a computer-readable storage medium and a robot. The method comprises: acquiring a first image and a second image, wherein the first image is an image to be subjected to semantic segmentation, and the second image is an image frame previous to the first image; performing optical-flow calculation on the first image and the second image, so as to obtain first optical-flow data; fusing the first optical-flow data with the first image, so as to obtain an optical-flow fusion image; and using a preset semantic segmentation model to perform semantic segmentation on the optical-flow fusion image, so as to obtain a semantic segmentation result corresponding to the first image. By means of the method, semantic segmentation can be performed in combination with optical-flow

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, 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))。

data of an image, such that deep features of the image can be mined, and semantic segmentation can be better performed on unknown objects, thereby effectively improving the accuracy of a semantic segmentation method.

(57) 摘要: 本申请属于机器人技术领域, 尤其涉及一种语义分割方法、装置、计算机可读存储介质及机器人。所述方法包括: 获取第一图像和第二图像; 其中, 所述第一图像为待进行语义分割的图像, 所述第二图像为所述第一图像的前一帧图像; 对所述第一图像和所述第二图像进行光流计算, 得到第一光流数据; 将所述第一光流数据与所述第一图像进行融合, 得到光流融合图像; 使用预设的语义分割模型对所述光流融合图像进行语义分割, 得到与所述第一图像对应的语义分割结果。通过上述方法, 可以结合图像的光流数据进行语义分割, 从而能够挖掘图像的深层特征, 能更好地对未知物体进行语义分割, 进而有效地提升了语义分割方法的准确率。

一种语义分割方法、装置、计算机可读存储介质及机器人

本申请要求于 2023 年 06 月 28 日在中国专利局提交的、申请号为 202310781685.7、发明名称为“一种语义分割方法、装置、计算机可读存储介质及机器人”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5 技术领域

本申请属于机器人技术领域，尤其涉及一种语义分割方法、装置、计算机可读存储介质及机器人。

背景技术

随着智能技术的发展，语义分割技术得以应用于自动驾驶、室内导航、虚拟现实和图像识别等多种场景。

在现有技术中，在利用语义分割模型对图像进行语义分割之前，通常需要通过语义分割训练集对语义分割模型进行训练。然而，当所在场景发生变化时，新的场景中可能会出现不存在于语义分割训练集中的未知物体，导致语义分割方法的准确率较低。

技术问题

有鉴于此，本申请实施例提供了一种语义分割方法、装置、计算机可读存储介质及机器人，以解决现有的语义分割方法在面对新场景时准确率较低的问题。

技术解决方案

本申请实施例的第一方面提供了一种语义分割方法，可以包括：

获取第一图像和第二图像；其中，所述第一图像为待进行语义分割的图像，所述第二图像为所述第一图像的前一帧图像；

对所述第一图像和所述第二图像进行光流计算，得到第一光流数据；

将所述第一光流数据与所述第一图像进行融合，得到光流融合图像；

使用预设的语义分割模型对所述光流融合图像进行语义分割，得到与所述第一图像对应的语义分割结果。

在第一方面的一种具体实现方式中，所述对所述第一图像和所述第二图像进行光流计算，得到第一光流数据，可以包括：

将所述第一图像和所述第二图像输入至预设的光流模型，并获取所述光流模型的中间输出结果；其中，所述中间输出结果为所述光流模型的最终输出结果之前的光流计算结果；

将所述中间输出结果作为所述第一光流数据。

在第一方面的一种具体实现方式中，所述将所述第一光流数据与所述第一图像进行融

合，得到光流融合图像，可以包括：

对所述第一光流数据进行插值，得到与所述第一图像的尺寸一致的插值光流结果；

将所述插值光流数据与所述第一图像进行融合，得到所述光流融合图像。

在第一方面的一种具体实现方式中，所述将所述插值光流数据与所述第一图像进行融

5 合，得到所述光流融合图像，可以包括：

将所述插值光流结果和所述第一图像在通道维度上进行拼接，得到所述光流融合图像。

在第一方面的一种具体实现方式中，在使用预设的语义分割模型对所述光流融合图像进行语义分割，得到与所述第一图像对应的语义分割结果之前，还可以包括：

获取第一样本图像、第二样本图像和与所述第一样本图像对应的语义分割标注；其中，

10 所述第二样本图像为所述第一样本图像的前一帧图像；

对所述第一样本图像和所述第二样本图像进行光流计算，得到第一光流样本数据；

将所述第一光流样本数据与所述第一样本图像进行融合，得到与所述第一样本图像对应的光流融合样本图像；

15 根据所述光流融合样本图像和所述语义分割标注对预设的深度学习模型进行训练，得到所述语义分割模型。

在第一方面的一种具体实现方式中，所述根据所述光流融合样本图像和所述语义分割标注对预设的深度学习模型进行训练，得到所述语义分割模型，可以包括：

将所述光流融合样本图像输入至所述深度学习模型，并获取所述深度学习模型的输出结果；

20 根据所述输出结果和所述语义分割标注计算所述深度学习模型的训练损失值；

根据所述训练损失值对所述深度学习模型进行模型参数调整，得到所述语义分割模型。

在第一方面的一种具体实现方式中，所述根据所述输出结果和所述语义分割标注计算所述深度学习模型的训练损失值，可以包括：

计算所述输出结果和所述语义分割标注之间的分割区域交并比；

25 根据所述分割区域交并比计算所述训练损失值。

本申请实施例的第二方面提供了一种语义分割装置，可以包括：

图像获取模块，用于获取第一图像和第二图像；其中，所述第一图像为待进行语义分割的图像，所述第二图像为所述第一图像的前一帧图像；

30 光流计算模块，用于对所述第一图像和所述第二图像进行光流计算，得到第一光流数据；

数据融合模块，用于将所述第一光流数据与所述第一图像进行融合，得到光流融合图

像；

语义分割模块，用于使用预设的语义分割模型对所述光流融合图像进行语义分割，得到与所述第一图像对应的语义分割结果。

在第二方面的一种具体实现方式中，所述光流计算模块可以包括：

- 5 结果获取单元，用于将所述第一图像和所述第二图像输入至预设的光流模型，并获取所述光流模型的中间输出结果；其中，所述中间输出结果为所述光流模型的最终输出结果之前的光流计算结果；

数据确定单元，用于将所述中间输出结果作为所述第一光流数据。

在第二方面的一种具体实现方式中，所述数据融合模块可以包括：

- 10 数据插值单元，用于对所述第一光流数据进行插值，得到与所述第一图像的尺寸一致的插值光流结果；

数据融合单元，用于将所述插值光流数据与所述第一图像进行融合，得到所述光流融合图像。

在第二方面的一种具体实现方式中，所述数据融合单元可以包括：

- 15 维度拼接子单元，用于将所述插值光流结果和所述第一图像在通道维度上进行拼接，得到所述光流融合图像。

在第二方面的一种具体实现方式中，所述语义分割装置还可以包括：

数据获取模块，用于获取第一样本图像、第二样本图像和与所述第一样本图像对应的语义分割标注；其中，所述第二样本图像为所述第一样本图像的前一帧图像；

- 20 数据计算模块，用于对所述第一样本图像和所述第二样本图像进行光流计算，得到第一光流样本数据；

图像融合模块，用于将所述第一光流样本数据与所述第一样本图像进行融合，得到与所述第一样本图像对应的光流融合样本图像；

- 25 模型训练模块，用于根据所述光流融合样本图像和所述语义分割标注对预设的深度学习模型进行训练，得到所述语义分割模型。

在第二方面的一种具体实现方式中，所述模型训练模块可以包括：

结果获取单元，用于将所述光流融合样本图像输入至所述深度学习模型，并获取所述深度学习模型的输出结果；

- 30 损失值计算单元，用于根据所述输出结果和所述语义分割标注计算所述深度学习模型的训练损失值；

参数调整单元，用于根据所述训练损失值对所述深度学习模型进行模型参数调整，得

到所述语义分割模型。

在第二方面的一种具体实现方式中，所述损失值计算单元可以包括：

第一计算子单元，用于计算所述输出结果和所述语义分割标注之间的分割区域交并比；

第二计算子单元，用于根据所述分割区域交并比计算所述训练损失值。

- 5 本申请实施例的第三方面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有计算机程序，所述计算机程序被处理器执行时实现上述任一种语义分割方法的步骤。

本申请实施例的第四方面提供了一种机器人，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现上述任一种语义分割方法的步骤。

- 10 本申请实施例的第五方面提供了一种计算机程序产品，当计算机程序产品在机器人上运行时，使得机器人执行上述任一种语义分割方法的步骤。

有益效果

- 本申请实施例与现有技术相比存在的有益效果是：本申请实施例获取第一图像和第二图像；其中，所述第一图像为待进行语义分割的图像，所述第二图像为所述第一图像的前一帧图像；对所述第一图像和所述第二图像进行光流计算，得到第一光流数据；将所述第
15 一光流数据与所述第一图像进行融合，得到光流融合图像；使用预设的语义分割模型对所述光流融合图像进行语义分割，得到与所述第一图像对应的语义分割结果。通过本申请实施例，可以结合图像的光流数据进行语义分割，从而能够挖掘图像的深层特征，能更好地对未知物体进行语义分割，进而有效地提升了语义分割方法的准确率。

20 附图说明

为了更清楚地说明本申请实施例中的技术方案，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，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其它的附图。

- 25 图 1 为语义分割模型训练过程的示意图；
图 2 为光流模型和语义分割模型训练过程的示意图；
图 3 为本申请实施例中一种语义分割方法的一个实施例流程图；
图 4 为本申请实施例中一种语义分割装置的一个实施例结构图；
图 5 为本申请实施例中一种机器人的示意框图。

30 本发明的实施方式

为使得本申请的发明目的、特征、优点能够更加的明显和易懂，下面将结合本申请实

施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，下面所描述的实施例仅仅是本申请一部分实施例，而非全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例，都属于本申请保护的范围。

5 应当理解，当在本说明书和所附权利要求书中使用时，术语“包括”指示所描述特征、整体、步骤、操作、元素和/或组件的存在，但并不排除一个或多个其它特征、整体、步骤、操作、元素、组件和/或其集合的存在或添加。

 还应当理解，在此本申请说明书中所使用的术语仅仅是出于描述特定实施例的目的而并不意在限制本申请。如在本申请说明书和所附权利要求书中所使用的那样，除非上下文清楚地指明其它情况，否则单数形式的“一”、“一个”及“该”意在包括复数形式。

10

 还应当进一步理解，在本申请说明书和所附权利要求书中使用的术语“和/或”是指相关联列出的项中的一个或多个的任何组合以及所有可能组合，并且包括这些组合。

 如在本说明书和所附权利要求书中所使用的那样，术语“如果”可以依据上下文被解释为“当...时”或“一旦”或“响应于确定”或“响应于检测到”。类似地，短语“如果确定”或“如果检测到[所描述条件或事件]”可以依据上下文被解释为意指“一旦确定”或“响应于确定”或“一旦检测到[所描述条件或事件]”或“响应于检测到[所描述条件或事件]”。

15

 另外，在本申请的描述中，术语“第一”、“第二”、“第三”等仅用于区分描述，而不能理解为指示或暗示相对重要性。

20 随着智能技术的发展，语义分割技术得以应用于自动驾驶、室内导航、虚拟现实和图像识别等多种场景。

 在现有技术中，在利用语义分割模型对图像进行语义分割之前，通常需要通过语义分割训练集对语义分割模型进行训练。然而，当所在场景发生变化时，新的场景中可能会出现不存在于语义分割训练集中的未知物体，导致语义分割方法的准确率较低。

25 有鉴于此，本申请实施例提供了一种语义分割方法、装置、计算机可读存储介质及机器人，以解决现有的语义分割方法在面对新场景时准确率较低的问题。通过本申请实施例，可以结合图像的光流数据进行语义分割，从而能够挖掘图像的深层特征，能更好地对未知物体进行语义分割，进而有效地提升了语义分割方法的准确率。

 需要说明的是，本申请方法的执行主体为机器人，可以包括但不限于扫地机、巡检机器人、送餐机器人、引导机器人等现有技术中任意一种常见的机器人。

30

 在本申请实施例中，可以通过预设的光流模型对图像进行光流计算，及通过预设的语

义分割模型对图像进行语义分割。因此，下面将首先对本申请实施例中的光流模型和语义分割模型进行介绍。

本申请实施例中的光流模型可以优选为轻量级的深度学习模型，通过光流模型，可以对图像进行光流估计，具体地，光流模型可以包括编码器（Encoder）和解码器（Decoder），其中，编码器主要可以用于进行金字塔多尺度特征提取，解码器主要可以用于进行光流估计。

可以理解的是，在利用光流模型进行光流估计之前，可以对原始的深度学习模型进行训练，得到光流模型。

具体地，可以获取连续的两帧样本图像，并利用现有技术中任意一种光流估计方法对样本图像的光流进行标注，得到预期光流数据（Ground Truth）；之后，可以将两帧样本图像作为原始的深度学习模型的输入，将对应的预期光流数据作为原始的深度学习模型的预期输出，对原始的深度学习模型进行训练。

可以理解的是，遍历获取到的预设数目的连续两帧样本图像，并重复执行上述训练过程，即可得到本申请实施例的光流模型。

在本申请实施例的一种具体实现方式中，为了保证原始的深度学习模型的训练效果，还可以利用预期光流数据对输出的预测光流数据进行监督，具体可以根据模型训练过程中输出的预测光流数据与预期光流数据，计算模型训练损失值，计算公式可以为：

$$\rho(x, \alpha, c) = \begin{cases} \frac{1}{2}(x/c)^2 & \alpha = 2 \\ \log\left(\frac{1}{2}(x/c)^2 + 1\right) & \alpha = 0 \\ 1 - \exp\left(-\frac{1}{2}(x/c)^2\right) & \alpha = -\infty \\ \frac{|\alpha - 2|}{\alpha} \left(\left(\frac{(x/c)^2}{|\alpha - 2|} + 1 \right)^{\alpha/2} - 1 \right) & \text{其它} \end{cases}$$

其中， x 为预测光流数据与预期光流数据之间的差值， α 和 c 为超参数，两个超参数的取值可以根据实际需要进行设置，本申请实施例对此不作具体限定。

可以理解的是，当训练损失值小于预设的损失值阈值时，或者，当模型训练的迭代次数大于预设的次数阈值时，则可以认为得到的光流模型较为准确，从而确定模型训练完成。

在本申请实施例中，还可以利用语义分割模型对图像进行语义分割，其中，语义分割模型可以为现有技术中任意一种可用于进行语义分割的深度学习模型。

可以理解的是，在利用语义分割模型对图像进行语义分割之前，可以对原始的深度学习

习模型进行训练，得到语义分割模型。具体地，可以构建语义分割训练样本集，并利用语义分割训练样本集对深度学习模型进行训练，此处，可以参阅图 1，语义分割模型的训练过程可以包括如下步骤：

步骤 S101、获取第一样本图像、第二样本图像和与第一样本图像对应的语义分割标注。

5 在本申请实施例中，可以使用预设的摄像采集装置进行视频采集，并可以获取视频中相邻的两帧图像，此处，可以获取视频中的第 t 帧图像和第 $t-1$ 帧图像，其中， t 为正整数，并将第 t 帧图像记为第一样本图像，将第 $t-1$ 帧图像记为第二样本图像。

另外，还可以获取与第一样本图像对应的语义分割标注，其中，图像的语义分割标注可以包括语义分割区域的像素个数和像素位置，即该图像中各个物体的像素个数和像素位置，具体地，可以使用现有技术中任何一种语义分割方法进行语义分割，分割出第一样本图像中的各个物体及各个物体分别对应的区域。

步骤 S102、对第一样本图像和第二样本图像进行光流计算，得到第一光流样本数据。

可以理解的是，图像的光流数据是图像的深层特征，可以用于描述图像中物体的运动，从而可以更容易分辨图像中的静态物体和动态物体，据此，可以有效区分图像的前景和背景，因此，当图像中存在未知物体时，引入图像的光流数据有助于进行未知物体的语义分割。

在本申请实施例中，具体可以使用光流模型对第一样本图像和第二样本图像进行光流计算。具体地，请参阅图 2，可以将第一样本图像和第二样本图像输入至光流模型，之后，第一样本图像与第二样本图像会经过编码器与解码器，此处，可以将解码器的中间输出结果（Half Size 输出结果）作为第一光流样本数据，其中，中间输出结果为光流模型最终输出结果之前的输出结果，具体可将解码器的中间层的输出结果作为中间输出结果，此处，中间输出结果的尺寸为最终输出结果的一半（Half Size）。

步骤 S103、将第一光流样本数据与第一样本图像进行融合，得到与第一样本图像对应的光流融合样本图像。

25 在本申请实施例中，可以对第一光流样本数据进行插值，得到与第一样本图像的尺寸一致的插值样本光流结果，具体地，可以利用现有技术中任何一种插值方法对第一光流数据进行插值，包括但不限于双线性插值、双三次插值和最近邻插值法等现有技术中任何一种常见的插值方法，优选地，此处可以使用双线性插值对第一光流数据进行插值。

之后，可以将插值样本光流结果和第一样本图像在通道维度上进行拼接（Concat），
30 得到光流融合样本图像，如下式所示：

$$Input_{Semantic} = Concat(Upsample(Flow), Frame_i)$$

其中, $Input_{Semantic}$ 为光流融合样本图像, $Flow$ 为第一光流样本数据, $Upsample(Flow)$ 为插值样本光流结果, $Frame_i$ 为第一样本图像, $Concat$ 为拼接操作。

5 步骤 S104、根据光流融合样本图像和语义分割标注对预设的深度学习模型进行训练, 得到语义分割模型。

在本申请实施例中, 具体可以将光流融合样本图像作为原始的深度学习模型的输入, 将光流融合样本图像对应的语义分割标注作为原始的深度学习模型的输出, 对原始的深度学习模型进行训练。

可以理解的是, 原始的深度学习模型的结构可以根据实际需要进行具体化、情景化设置, 本申请对此不作限定。在本申请实施例中, 原始深度学习模型优选可以包括卷积层和池化层等, 并可以进行金字塔多尺度特征提取、池化、拼接等常见的操作。

在本申请实施例的一种具体实现方式中, 还可以根据原始的深度学习模型的输出结果和语义分割标注计算训练损失值, 此处, 可以计算输出结果和语义分割标注之间的分割区域交并比 (Intersection-over Union, IoU), 其中, 交并比可用于表示两个结果的分割区域之间的重合度, 计算公式如下:

$$dice = \frac{2|X \cap Y|}{|X| + |Y|}$$

具体地, $dice$ 为输出结果和语义分割标注之间的分割区域交并比, X 为深度学习模型的输出结果, 即输出的语义分割区域的像素个数和像素位置, Y 为语义分割标注, 即第一样本图像的语义分割区域的像素个数和像素位置, $|X \cap Y|$ 为输出的语义分割区域与第一样本图像的语义分割区域之间的交集, $|X| + |Y|$ 为输出的语义分割区域与第一样本图像的语义分割区域之间的并集。

之后, 可以根据该分割区域交并比计算训练损失值 L_{dice} , 计算公式可如下:

$$L_{dice} = 1 - \frac{2|X \cap Y|}{|X| + |Y|}$$

在得到训练损失值后, 可以根据训练损失值对深度学习模型的模型参数进行调整。例如, 可以根据反向更新的方式对深度学习模型进行参数调整。

可以理解的是, 当训练损失值小于预设的损失值阈值时, 或者, 当模型训练的迭代次数大于预设的次数阈值时, 可以认为得到的语义分割模型较为准确, 此时, 可以使用该语

义分割模型对图像进行语义分割，请参阅图 3，本申请实施例中一种语义分割方法的一个实施例可以包括：

步骤 S301、获取第一图像和第二图像。

在本申请实施例中，优选可以使用单目相机对待进行语义分割的区域进行视频采集，
5 并将采集到的视频中第 t 帧图像作为第一图像，将第 $t-1$ 帧图像作为第二图像。

例如，本申请方法具体可应用于机器人的导航过程，则待进行语义分割的区域可以为机器人的行驶区域，此时可以利用安装于机器人的单目相机对行驶区域进行视频采集，并将采集到的视频中的第 t 帧图像作为第一图像，将视频中的第 $t-1$ 帧图像作为第二图像。

步骤 S302、对第一图像和第二图像进行光流计算，得到第一光流数据。

10 在本申请实施例中，具体可以将第一图像和第二图像输入至训练得到的光流模型，并可以获取光流模型的中间输出结果，将该中间输出结果作为第一光流数据。

步骤 S303、将第一光流数据与第一图像进行融合，得到光流融合图像。

在本申请实施例中，为了使得第一光流数据能顺利与第一图像进行融合，可以对第一光流数据进行插值，得到与第一图像尺寸一致的插值光流结果，之后，可以将插值光流数
15 据与第一图像在通道维度上进行拼接，得到光流融合图像。

可以理解的是，步骤 S303 的具体内容可参考步骤 S103 的相关描述，此处不再赘述。

步骤 S304、使用预设的语义分割模型对光流融合图像进行语义分割，得到与第一图像对应的语义分割结果。

在本申请实施例中，可以使用训练好的语义分割模型对光流融合图像进行语义分割，
20 具体地，可以将光流融合图像作为语义分割模型的输入，即可得到与第一图像对应的语义分割结果。其中，语义分割结果可以包括已知物体和未知物体的相关信息。

在本申请实施例的一种具体实现方式中，可以利用与第一图像对应的语义分割结果进行导航、探索及运动控制等相关应用。例如，本申请方法可以应用于导航的具体应用场景，第一图像可以为采集到的行驶区域的图像，则可以利用语义分割模型分割出行驶区域中的
25 可行驶区域、障碍物、行人等物体，并据此进行导航。

综上所述，本申请实施例获取第一图像和第二图像；其中，所述第一图像为待进行语义分割的图像，所述第二图像为所述第一图像的前一帧图像；对所述第一图像和所述第二图像进行光流计算，得到第一光流数据；将所述第一光流数据与所述第一图像进行融合，得到光流融合图像；使用预设的语义分割模型对所述光流融合图像进行语义分割，得到与
30 所述第一图像对应的语义分割结果。通过本申请实施例，可以结合图像的光流数据进行语义分割，从而能够挖掘图像的深层特征，能更好地对未知物体进行语义分割，进而有效地

提升了语义分割方法的准确率。

应理解，上述实施例中各步骤的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能和内在逻辑确定，而不对本申请实施例的实施过程构成任何限定。

5 对应于上文实施例所述的一种语义分割方法，图4示出了本申请实施例提供的一种语义分割装置的一个实施例结构图。

本实施例中，一种语义分割装置可以包括：

图像获取模块401，用于获取第一图像和第二图像；其中，所述第一图像为待进行语义分割的图像，所述第二图像为所述第一图像的前一帧图像；

10 光流计算模块402，用于对所述第一图像和所述第二图像进行光流计算，得到第一光流数据；

数据融合模块403，用于将所述第一光流数据与所述第一图像进行融合，得到光流融合图像；

15 语义分割模块404，用于使用预设的语义分割模型对所述光流融合图像进行语义分割，得到与所述第一图像对应的语义分割结果。

在本申请实施例的一种具体实现方式中，所述光流计算模块可以包括：

结果获取单元，用于将所述第一图像和所述第二图像输入至预设的光流模型，并获取所述光流模型的中间输出结果；其中，所述中间输出结果为所述光流模型的最终输出结果之前的光流计算结果；

20 数据确定单元，用于将所述中间输出结果作为所述第一光流数据。

在本申请实施例的一种具体实现方式中，所述数据融合模块可以包括：

数据插值单元，用于对所述第一光流数据进行插值，得到与所述第一图像的尺寸一致的插值光流结果；

25 数据融合单元，用于将所述插值光流数据与所述第一图像进行融合，得到所述光流融合图像。

在本申请实施例的一种具体实现方式中，所述数据融合单元可以包括：

维度拼接子单元，用于将所述插值光流结果和所述第一图像在通道维度上进行拼接，得到所述光流融合图像。

在本申请实施例的一种具体实现方式中，所述语义分割装置还可以包括：

30 数据获取模块，用于获取第一样本图像、第二样本图像和与所述第一样本图像对应的语义分割标注；其中，所述第二样本图像为所述第一样本图像的前一帧图像；

数据计算模块，用于对所述第一样本图像和所述第二样本图像进行光流计算，得到第一光流样本数据；

图像融合模块，用于将所述第一光流样本数据与所述第一样本图像进行融合，得到与所述第一样本图像对应的光流融合样本图像；

- 5 模型训练模块，用于根据所述光流融合样本图像和所述语义分割标注对预设的深度学习模型进行训练，得到所述语义分割模型。

在本申请实施例的一种具体实现方式中，所述模型训练模块可以包括：

结果获取单元，用于将所述光流融合样本图像输入至所述深度学习模型，并获取所述深度学习模型的输出结果；

- 10 损失值计算单元，用于根据所述输出结果和所述语义分割标注计算所述深度学习模型的训练损失值；

参数调整单元，用于根据所述训练损失值对所述深度学习模型进行模型参数调整，得到所述语义分割模型。

在本申请实施例的一种具体实现方式中，所述损失值计算单元可以包括：

- 15 第一计算子单元，用于计算所述输出结果和所述语义分割标注之间的分割区域交并比；
第二计算子单元，用于根据所述分割区域交并比计算所述训练损失值。

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，模块和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

- 在上述实施例中，对各个实施例的描述都各有侧重，某个实施例中没有详述或记载的
20 部分，可以参见其它实施例的相关描述。

图5示出了本申请实施例提供的一种机器人的示意框图，为了便于说明，仅示出了与本申请实施例相关的部分。

- 如图5所示，该实施例的机器人5包括：处理器50、存储器51以及存储在所述存储器51中并可在所述处理器50上运行的计算机程序52。所述处理器50执行所述计算机程序52时实现上述各个语义分割方法实施例中的步骤，例如图3所示的步骤S301至步骤S304。或者，所述处理器50执行所述计算机程序52时实现上述各装置实施例中各模块/单元的功能，例如图4所示模块401至模块404的功能。

- 示例性的，所述计算机程序52可以被分割成一个或多个模块/单元，所述一个或者多个模块/单元被存储在所述存储器51中，并由所述处理器50执行，以完成本申请。所述一个或多个模块/单元可以是能够完成特定功能的一系列计算机程序指令段，该指令段用于描
30

述所述计算机程序 52 在所述机器人 5 中的执行过程。

本领域技术人员可以理解,图 5 仅仅是机器人 5 的示例,并不构成对机器人 5 的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件,例如所述机器人 5 还可以包括输入输出设备、网络接入设备、总线等。

5 所述处理器 50 可以是中央处理单元 (Central Processing Unit, CPU), 还可以是其它通用处理器、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) 或者其它可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

10 所述存储器 51 可以是所述机器人 5 的内部存储单元,例如机器人 5 的硬盘或内存。所述存储器 51 也可以是所述机器人 5 的外部存储设备,例如所述机器人 5 上配备的插接式硬盘,智能存储卡 (Smart Media Card, SMC),安全数字 (Secure Digital, SD) 卡,闪存卡 (Flash Card) 等。进一步地,所述存储器 51 还可以既包括所述机器人 5 的内部存储单元也包括外部存储设备。所述存储器 51 用于存储所述计算机程序以及所述机器人 5 所需的
15 其它程序和数据。所述存储器 51 还可以用于暂时地存储已经输出或者将要输出的数据。

 所属领域的技术人员可以清楚地了解到,为了描述的方便和简洁,仅以上述各功能单元、模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能单元、模块完成,即将所述装置的内部结构划分成不同的功能单元或模块,以完成以上描述的全部或者部分功能。实施例中的各功能单元、模块可以集成在一个处理单元中,也
20 可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中,上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。另外,各功能单元、模块的具体名称也只是为了便于相互区分,并不用于限制本申请的保护范围。上述系统中单元、模块的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

25 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中没有详述或记载的部分,可以参见其它实施例的相关描述。

 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以
30 对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本申请的范围。

在本申请所提供的实施例中，应该理解到，所揭露的装置/机器人和方法，可以通过其它的方式实现。例如，以上所描述的装置/机器人实施例仅仅是示意性的，例如，所述模块或单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通讯连接可以是通过一些接口，装置或单元的间接耦合或通讯连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

所述集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读存储介质中。基于这样的理解，本申请实现上述实施例方法中的全部或部分流程，也可以通过计算机程序来指令相关的硬件来完成，所述的计算机程序可存储于一计算机可读存储介质中，该计算机程序在被处理器执行时，可实现上述各个方法实施例的步骤。其中，所述计算机程序包括计算机程序代码，所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读存储介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U 盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、电载波信号、电信信号以及软件分发介质等。需要说明的是，所述计算机可读存储介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减，例如在某些司法管辖区，根据立法和专利实践，计算机可读存储介质不包括电载波信号和电信信号。

以上所述实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围，均应包含在本申请的保护范围之内。

权利要求书

1. 一种语义分割方法，其特征在于，包括：

获取第一图像和第二图像；其中，所述第一图像为待进行语义分割的图像，所述第二图像为所述第一图像的前一帧图像；

5 对所述第一图像和所述第二图像进行光流计算，得到第一光流数据；

将所述第一光流数据与所述第一图像进行融合，得到光流融合图像；

使用预设的语义分割模型对所述光流融合图像进行语义分割，得到与所述第一图像对应的语义分割结果。

2. 根据权利要求1所述的语义分割方法，其特征在于，所述对所述第一图像和所述第二图像进行光流计算，得到第一光流数据，包括：

将所述第一图像和所述第二图像输入至预设的光流模型，并获取所述光流模型的中间输出结果；其中，所述中间输出结果为所述光流模型的最终输出结果之前的光流计算结果；

将所述中间输出结果作为所述第一光流数据。

3. 根据权利要求1所述的语义分割方法，其特征在于，所述将所述第一光流数据与所述第一图像进行融合，得到光流融合图像，包括：

对所述第一光流数据进行插值，得到与所述第一图像的尺寸一致的插值光流结果；

将所述插值光流数据与所述第一图像进行融合，得到所述光流融合图像。

4. 根据权利要求3所述的语义分割方法，其特征在于，所述将所述插值光流数据与所述第一图像进行融合，得到所述光流融合图像，包括：

20 将所述插值光流结果和所述第一图像在通道维度上进行拼接，得到所述光流融合图像。

5. 根据权利要求1至4中任一项所述的语义分割方法，其特征在于，在使用预设的语义分割模型对所述光流融合图像进行语义分割，得到与所述第一图像对应的语义分割结果之前，还包括：

25 获取第一样本图像、第二样本图像和与所述第一样本图像对应的语义分割标注；其中，所述第二样本图像为所述第一样本图像的前一帧图像；

对所述第一样本图像和所述第二样本图像进行光流计算，得到第一光流样本数据；

将所述第一光流样本数据与所述第一样本图像进行融合，得到与所述第一样本图像对应的光流融合样本图像；

30 根据所述光流融合样本图像和所述语义分割标注对预设的深度学习模型进行训练，得到所述语义分割模型。

6. 根据权利要求5所述的语义分割方法，其特征在于，所述根据所述光流融合样本图

像和所述语义分割标注对预设的深度学习模型进行训练，得到所述语义分割模型，包括：

将所述光流融合样本图像输入至所述深度学习模型，并获取所述深度学习模型的输出结果；

根据所述输出结果和所述语义分割标注计算所述深度学习模型的训练损失值；

5 根据所述训练损失值对所述深度学习模型进行模型参数调整，得到所述语义分割模型。

7. 根据权利要求 6 所述的语义分割方法，其特征在于，所述根据所述输出结果和所述语义分割标注计算所述深度学习模型的训练损失值，包括：

计算所述输出结果和所述语义分割标注之间的分割区域交并比；

根据所述分割区域交并比计算所述训练损失值。

10 8. 一种语义分割装置，其特征在于，包括：

图像获取模块，用于获取第一图像和第二图像；其中，所述第一图像为待进行语义分割的图像，所述第二图像为所述第一图像的前一帧图像；

光流计算模块，用于对所述第一图像和所述第二图像进行光流计算，得到第一光流数据；

15 数据融合模块，用于将所述第一光流数据与所述第一图像进行融合，得到光流融合图像；

语义分割模块，用于使用预设的语义分割模型对所述光流融合图像进行语义分割，得到与所述第一图像对应的语义分割结果。

9. 一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，其特征在于，所述计算机程序被处理器执行时实现如权利要求 1 至 7 中任一项所述的语义分割方法的步骤。

10. 一种机器人，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序，其特征在于，所述处理器执行所述计算机程序时实现如权利要求 1 至 7 中任一项所述的语义分割方法的步骤。

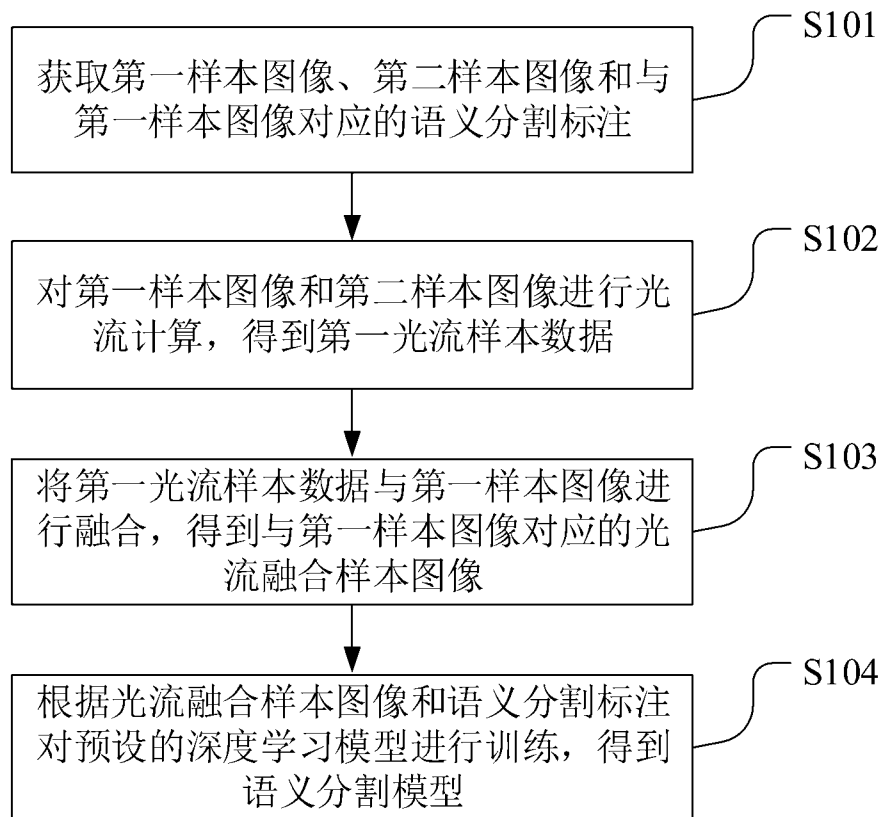


图 1

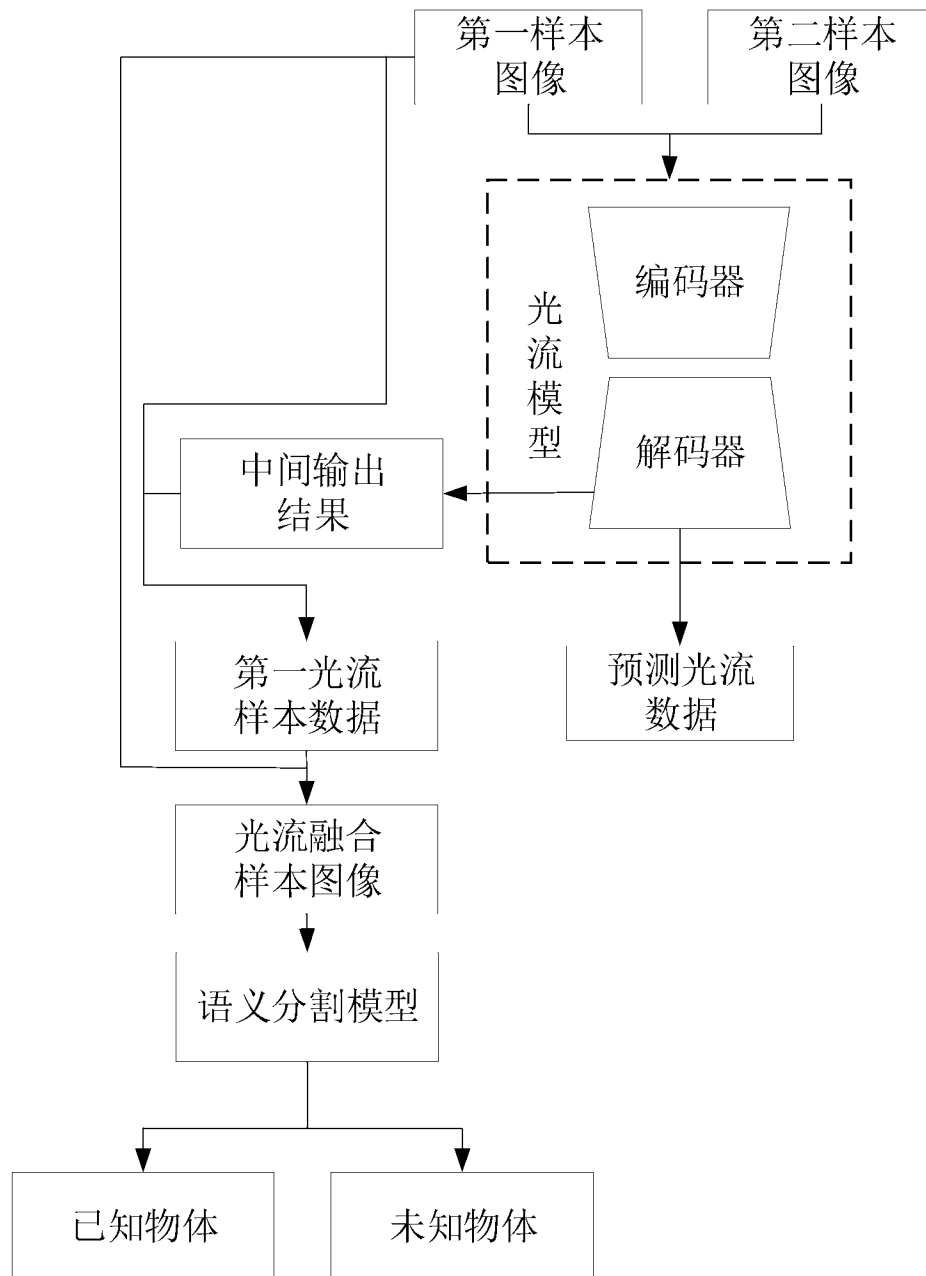


图 2

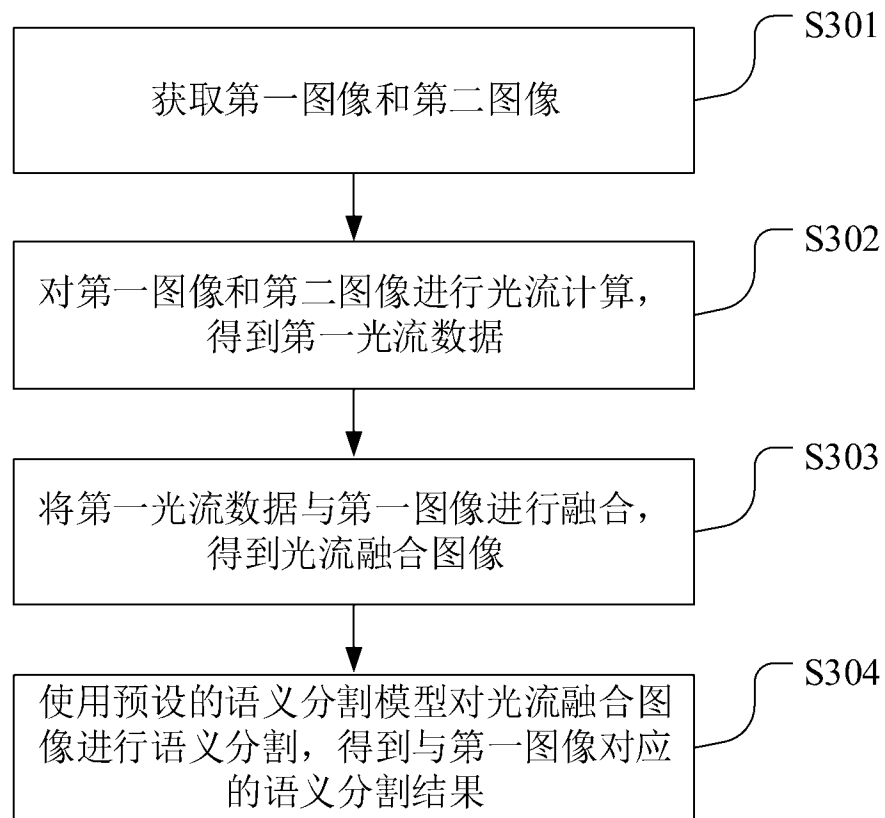


图 3

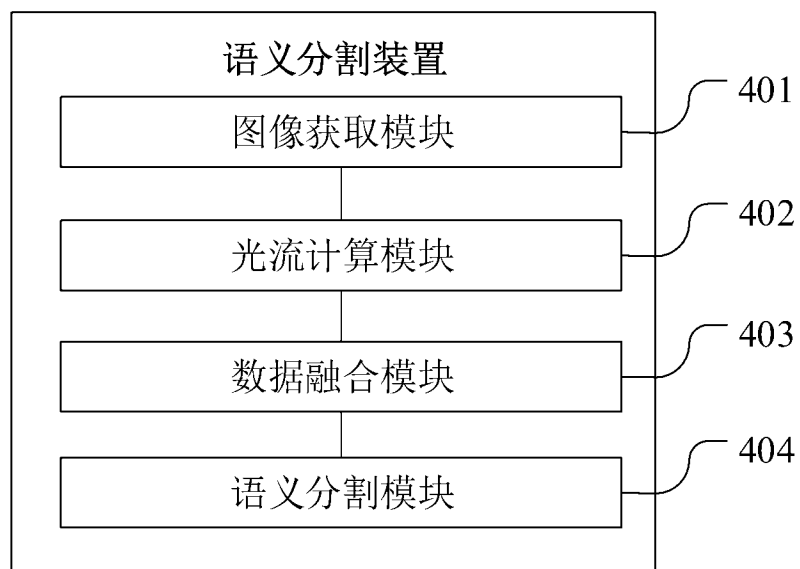


图 4

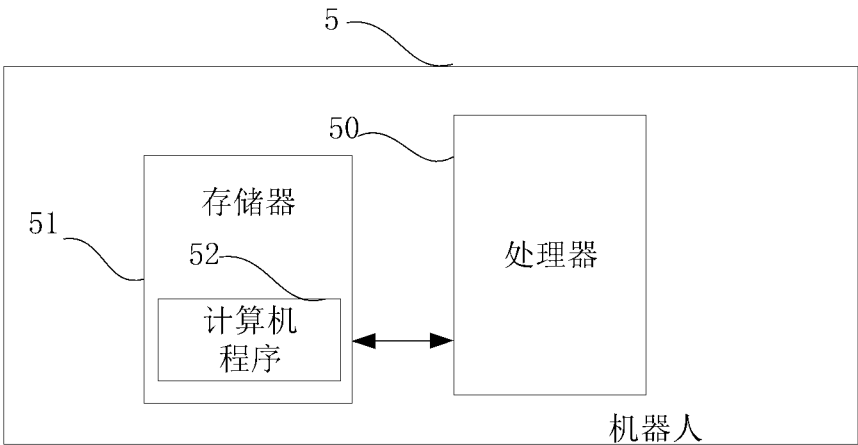


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/141742

A. CLASSIFICATION OF SUBJECT MATTER

G06V 10/26(2022.01)i; G06V 20/70(2022.01)i; G06V 10/774(2022.01)i; G06V 10/80(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

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; VEN; USTXT; WOTXT; EPTXT: 语义, 分割, 光流, 融合, 模型, 插值, 拼接, 标注, semantic, segmentation, optical flow, fusion, model, interpolation, stitching, label

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 116883656 A (UBTECH ROBOTICS CORP.) 13 October 2023 (2023-10-13) claims 1-10, and description, paragraphs [0005]-[0153]	1-10
Y	CN 112862839 A (TSINGHUA UNIVERSITY) 28 May 2021 (2021-05-28) description, paragraphs [0007]-[0066]	1-10
Y	CN 114283352 A (BEIJING KUANGSHI TECHNOLOGY CO., LTD. et al.) 05 April 2022 (2022-04-05) description, paragraphs [0006]-[0034]	1-10
A	US 2021353393 A1 (RETRACE LABS) 18 November 2021 (2021-11-18) entire document	1-10



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

25 February 2024

Date of mailing of the international search report

07 March 2024

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/CN2023/141742

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	116883656	A	13 October 2023	None			
CN	112862839	A	28 May 2021	CN	112862839	B	23 December 2022
CN	114283352	A	05 April 2022	None			
US	2021353393	A1	18 November 2021	US	11357604	B2	14 June 2022

A. 主题的分类

G06V 10/26(2022.01)i; G06V 20/70(2022.01)i; G06V 10/774(2022.01)i; G06V 10/80(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

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))

CNTEXT;CNABS;CNKI;VEN;USTXT;WOTXT;EPTXT: 语义, 分割, 光流, 融合, 模型, 插值, 拼接, 标注, semantic, segmentation, optical flow, fusion, model, interpolation, stitching, label

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 116883656 A (深圳市优必选科技股份有限公司) 2023年10月13日 (2023 - 10 - 13) 权利要求第1-10项, 说明书第[0005]-[0153]段	1-10
Y	CN 112862839 A (清华大学) 2021年5月28日 (2021 - 05 - 28) 说明书第[0007]-[0066]段	1-10
Y	CN 114283352 A (北京旷视科技有限公司 等) 2022年4月5日 (2022 - 04 - 05) 说明书第[0006]-[0034]段	1-10
A	US 2021353393 A1 (RETRACE LABS) 2021年11月18日 (2021 - 11 - 18) 全文	1-10

☐ 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

2024年2月25日

国际检索报告邮寄日期

2024年3月7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

徐晓艳

电话号码 (+86) 0512-88995723

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2023/141742

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	116883656	A	2023年10月13日	无	
CN	112862839	A	2021年5月28日	CN	112862839 B 2022年12月23日
CN	114283352	A	2022年4月5日	无	
US	2021353393	A1	2021年11月18日	US	11357604 B2 2022年6月14日