

(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 5 月 2 日 (02.05.2024)



(10) 国际公布号
WO 2024/087331 A1

(51) 国际专利分类号:

B25J 15/08 (2006.01) **G06V 10/80** (2022.01)
B25J 9/16 (2006.01) **G06V 10/764** (2022.01)
B25J 13/08 (2006.01) **G06V 10/82** (2022.01)
G06N 3/04 (2023.01)

(21) 国际申请号: PCT/CN2022/137001

(22) 国际申请日: 2022 年 12 月 6 日 (06.12.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:

202211303570.9 2022年10月24日 (24.10.2022) CN

(71) 申请人: 深圳先进技术研究院 (SHENZHEN INSTITUTE OF ADVANCED TECHNOLOGY) [CN/CN]; 中国广东省深圳市南山区西丽大学城学苑大道1068号, Guangdong 518055 (CN)。

(72) 发明人: 易正琨 (YI, Zhengkun); 中国广东省深圳市南山区西丽大学城学苑大道1068号, Guangdong 518055 (CN)。 柳程亮 (LIU, Chengliang); 中国广

东省深圳市南山区西丽大学城学苑大道 1068 号, Guangdong 518055 (CN)。 吴新宇 (WU, Xinyu); 中国广东省深圳市南山区西丽大学城学苑大道1068号, Guangdong 518055 (CN)。 崔允端 (CUI, Yunduan); 中国广东省深圳市南山区西丽大学城学苑大道1068号, Guangdong 518055 (CN)。 谢祥 (XIE, Xiang); 中国广东省深圳市南山区西丽大学城学苑大道1068号, Guangdong 518055 (CN)。 田琼 (TIAN, Qiong); 中国广东省深圳市南山区西丽大学城学苑大道1068号, Guangdong 518055 (CN)。

(74) 代理人: 北京中巡通大知识产权代理有限公司 (BEIJING ZHONG XUN TONG DA INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国北京市西城区西直门外大街1号院2号楼7层7C1, Beijing 100044 (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: ROBOTIC GRASPING PREDICTION METHOD BASED ON TRIPLET CONTRASTIVE NETWORK

(54) 发明名称: 一种基于三元组对比网络的机器人抓取预测方法

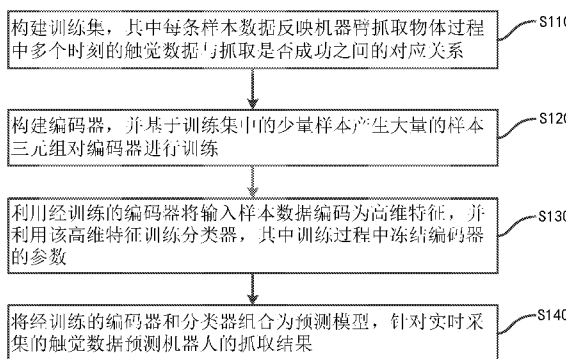


图 1

- S110 Construct a training set, in which each piece of sample data reflects the corresponding relationship between the success and failure of grasping and tactile data at a plurality of moments during the process of a robotic arm grasping an object
- S120 Construct encoders and, on the basis of a small number of samples in the training set, generate a large number of sample triplets to train the encoders
- S130 Use the trained encoders to encode input sample data into high-dimensional features, and use the high-dimensional features to train a classifier, parameters of the encoders being frozen during the process of training
- S140 Combine the trained encoders and classifier into a prediction model to predict a grasping result of a robot in respect of the tactile data collected in real time

(57) Abstract: Disclosed in the present invention is a robotic grasping prediction method based on a triplet contrastive network. The method comprises: constructing a training set, in which each sample reflects the corresponding relationship between classification labels and tactile data at a plurality of moments during the process of a robotic arm grasping an object; training encoders on the basis of a set loss function and, during the process of training, extracting at random from the training set two samples having the same label and taking same as an anchor sample and a positive sample, and extracting at random a sample having an opposite label and taking same as a negative sample; inputting into the corresponding encoders the sample data to be encoded so as to obtain the high-dimensional feature of the anchor sample, the high-dimensional feature of the positive sample and the high-dimensional feature of the negative sample; freezing optimized parameters of the encoders, using the trained encoders to encode the input sample data into high-dimensional features, and then inputting the high-dimensional features into a classifier for training; and using the trained encoders and classifier to predict a grasping result with regard to the real-time tactile data. The present invention improves the accuracy of robotic grasping prediction.

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))。

(57) 摘要: 本发明公开一种基于三元组对比网络的机器人抓取预测方法。该方法包括: 构建训练集, 其中每条样本反映机器臂抓取物体过程中多个时刻的触觉数据与分类标签之间的对应关系; 基于设定的损失函数训练编码器, 在训练过程中, 在训练集随机抽取两个标签一致的样本分别作为锚样本和正样本, 并随机抽取一个标签相反的样本作为负样本, 将样本数据分别输入对应编码器中编码得到锚样本的高维特征、正样本的高维特征以及负样本的高维特征; 冻结编码器的优化参数, 并利用经训练的编码器将输入样本数据编码为高维特征, 进而将该高维特征输入分类器进行训练; 利用经训练的编码器和分类器针对实时的触觉数据预测抓取结果。本发明提升了机器人抓取预测的准确率。

一种基于三元组对比网络的机器人抓取预测方法

技术领域

本发明涉及机器人技术领域，更具体地，涉及一种基于三元组对比网络的机器人抓取预测方法。

背景技术

随着时代的发展，智能机器人技术已经在全球众多行业内广泛应用，机器人的应用场景越来越丰富。例如，从汽车装配、金属加工到 3C 制造、家电生产、食品包装以及餐饮服务等行业均存在机器人应用。在这些应用场景中，机器人抓取是一个常见并且有待深入研究优化的领域。在执行抓取之前，根据现有信息准确预测本次抓取的成功率对于进一步生成高质量稳定的抓取大有裨益。

在现有技术中，专利申请 CN114700947A 提出了一种视触融合的抓取滑动概率预测方法，该方法在机器手上同时安装有视觉相机和触觉传感器，用于获得两个模态的信息。视觉与触觉感知融合算法利用采集的视觉信息获取目标物体的外部轮廓信息，触觉信息获取目标物体的表面形貌与质地软硬等信息，将采集的信息预处理后利用卷积神经网络提取视觉信息的特征和触觉信息的特征，两个模态特征信息融合之后获得的新的特征用于预测物体抓取过程中的滑动概率。

专利申请 CN114083535A 提供了一种机器手抓取姿势质量的物理度量方法及装置。该方法包括：确定机器手的候选抓取姿势与物体之间接触点的平坦度分数；确定所述候选抓取姿势夹持所述物体的重心分数；基于所述平坦度分数和所述重心分数评估候选抓取姿势的质量。该方案提供的机器手抓取姿势质量的物理度量方法，基于人类在生活中拿起物体时倾向于接触物体上更平坦的部位以及习惯于接触物体的中心部位的特性，通过两种物理度量分数分别从物体自身平坦度和重力分析上评价机器手抓取姿势的质量。

上述现有技术虽然从不同的角度提出了能够预测抓取是否成功的方案，但就训练集的数据量而言，一些现有方案基于数据驱动，需要大量抓取数据

作为训练集支撑，训练代价高昂。从特征提取方法上来看，一些现有方案训练过程未考虑样本与样本之间的关系，没能充分发掘和利用样本信息，或者自创构建手工特征，局限性强，泛化性弱，可信度不高。

综上，人们在抓握物体时可以感受预测当前的力度是否足以稳定的抓取物品，而近期研究表明，通过机器人抓取物品提升阶段之前的触觉信息，也可以有效预测机器人抓取是否成功。然而，当前的机器人抓取预测算法在训练过程中忽略了样本与样本之间的关系，仅仅通过单纯的端到端的方式来训练。此外，这些方法通常基于数据驱动，在可供训练的抓取样本数据较少的情况下难以有效训练模型。

发明内容

本发明的目的是克服上述现有技术的缺陷，提供一种基于三元组对比网络的机器人抓取预测方法。该方法包括以下步骤：

构建训练集，其中每条样本数据反映机器臂抓取物体过程中多个时刻的触觉数据与分类标签之间的对应关系，该分类标签用于指示抓取是否成功；

基于设定的损失函数训练编码器，获得优化参数，其中在训练过程中，在所述训练集随机抽取两个标签一致的样本分别作为锚样本和正样本，并随机抽取一个标签相反的样本作为负样本；将样本数据分别输入对应编码器中编码得到锚样本的高维特征、正样本的高维特征以及负样本的高维特征；

冻结所述编码器的优化参数，并利用经训练的编码器将输入样本数据编码为高维特征，进而将该高维特征输入分类器进行训练；

利用经训练的编码器和经训练的分类器针对实时采集的触觉数据预测抓取结果。

与现有技术相比，本发明的优点在于，所提出的基于三元组对比网络的机器人抓取预测方法，利用深度神经网络，通过将样本相互对比来训练编码器，提高了样本利用率，能够从更少的样本中获得更佳的网络参数。并且，本发明在网络中加入了自注意力机制，该机制使网络能更有效的捕捉特征的内部相关性，从而进一步提升了网络预测准确率。

通过以下参照附图对本发明的示例性实施例的详细描述，本发明的其

它特征及其优点将会变得清楚。

附图说明

被结合在说明书中并构成说明书的一部分的附图示出了本发明的实施例，并且连同其说明一起用于解释本发明的原理。

图 1 是根据本发明一个实施例的基于三元组对比网络的机器人抓取预测方法的流程图；

图 2 是根据本发明一个实施例的编码器网络的示意图；

图 3 是根据本发明一个实施例的编码器对比学习训练的示意图；

图 4 是根据本发明一个实施例的分类器训练示意图；

图 5 是根据本发明一个实施例的基于三元组对比网络的机器人抓取预测方法的过程示意图。

具体实施方式

现在将参照附图来详细描述本发明的各种示例性实施例。应注意到：除非另外具体说明，否则在这些实施例中阐述的部件和步骤的相对布置、数字表达式和数值不限制本发明的范围。

以下对至少一个示例性实施例的描述实际上仅仅是说明性的，决不作为对本发明及其应用或使用的任何限制。

对于相关领域普通技术人员已知的技术、方法和设备可能不作详细讨论，但在适当情况下，所述技术、方法和设备应当被视为说明书的一部分。

在这里示出和讨论的所有例子中，任何具体值应被解释为仅仅是示例性的，而不是作为限制。因此，示例性实施例的其它例子可以具有不同的值。

应注意到：相似的标号和字母在下面的附图中表示类似项，因此，一旦某一项在一个附图中被定义，则在随后的附图中不需要对其进行进一步讨论。

本发明主要涉及一种基于三元组对比网络的机器人抓取预测方法，属于机器人技术领域。其技术方案内容如下：

参见图 1 和图 5 所示，所提供的基于三元组对比网络的机器人抓取预

测方法包括以下步骤：

步骤 S110，构建训练集，其中每条样本数据反映机器臂抓取物体过程中多个时刻的触觉数据与抓取是否成功之间的对应关系。

在实际应用中，可以通过自主采集或在线下载公开数据集，以获取基于触觉或包含触觉信息的机器人抓取数据集。

例如，采用 Calandra 公开数据集，数据集中使用一个在两指上分别安装有 GELSIGHT 触觉传感器的 Weiss WSG-50 机械臂采集触觉数据，在抓取工作台前置放一个 Microsoft Kinect 2 深度相机采集视觉数据，抓取 106 种不同物品共 9296 次，数据集采集过程中，机器臂置于初始位置时，标记为时刻 T_a ，机器臂抓住目标物品但尚未提升时，标记为时刻 T_b ，抓住物品在空中悬停 2 秒后的时刻为 T_c ，针对这三个时刻获取对应的视触觉数据。

进一步地，对数据集进行预处理来增强数据，包括随机水平翻转，随机裁剪，归一化处理等，并将其划分为训练集和测试集。

例如，仅使用 T_a 、 T_b 时刻的触觉数据。作为数据增强，将触觉数据图像随机翻转（来自两指的触觉数据同时翻转或不反转），并将 256×256 的触觉数据随机裁剪为 224×224 。

在一个实施例中，为了提高训练的效率并降低对样本量的依赖，基于小样本场景划分训练集和测试集。例如，从抓取物品类别数目的小样本角度，从数据集中随机选取 10 种抓取物品的抓取数据构成训练集。从抓取数据的小样本角度，从每个物品的抓取数据中随机选取 20 个数据作为训练集样本。于是可以得到一个仅有 200 个样本的训练集，此外，从剩余未被训练的物品抓取数据中随机选取 1000 次抓取数据作为测试集。

综上，所构建的训练集样本量较少，每条样本数据反映机器臂抓取物体过程中多个时刻的触觉数据（或触觉以及视觉数据）与分类标签之间的对应关系，分类标签用于指示抓取是否成功。

步骤 S120，构建编码器，并基于训练集中的少量样本产生大量的样本三元组对编码器进行训练。

具体地，步骤 S120 包括以下子步骤：

步骤 S121，构建编码器

编码器训练框架适应不同数目的触觉传感器情景。对于只有单个触觉传感器的情况，将传感器在 T_a 、 T_b 时刻采集到的触觉数据分别表示为 G_a 、 G_b ，将 G_a 与 G_b 作差以突出两时刻间触觉信息的变化部分，并将 G_a 与 $G_b - G_a$ 在通道方向叠加为一个六通道图像输入预训练的去掉最后一层的 Resnet-50 骨干网络以提取高维特征。对于由多个触觉传感器的情况，将不同传感器的数据以上述形式进行处理后，将不同触觉传感器数据提取的高维特征拼接得到高维特征。然后，将其输入自注意力模块得到融合高维特征，双触觉传感器情景下编码器示意图参见图 2，该编码器包含两个骨干网络，每个骨干网络对应一个传感器的触觉数据。应理解的是，在多传感器且各传感器型号相同、与目标物体接触情况近似时，各骨干网络可以共享参数以节约模型空间大小，或不共享参数以拟合更复杂的映射网络获得更优的性能。

应理解的是，所采用的骨干网络，可使用除了 Resnet-50 以外的网络代替，如 VGG、AlexNet 等。此外，输入的基于视觉的触觉模态信息，理论上也可以由其他模态数据代替，如视觉拍照信息等。

步骤 S122，训练编码器

参见图 3 所示，在训练过程中，在训练集随机抽取两个标签一致（即均抓取成功或失败）的样本分别作为锚样本和正样本，再随机抽取一个标签相反的样本作为负样本，将样本数据分别输入编码器中编码得到锚样本、正样本、负样本的高维特征分别表示为 f_a 、 f_p 、 f_n ，然后，应用三元对比损失函数（Triplet Contrastive Loss Function）作为评价指标，将其公式化表示如下：

$$L_T = \left\| f_a - f_p \right\|_2^2 - \left\| f_a - f_n \right\|_2^2 + 1 \quad (1)$$

其中 L_T 表示三元对比损失。通过将 L_T 最小化，编码器将使同类标签样本编码高维特征之间的欧氏距离尽量小，异类样本编码高维特征之间的欧氏距离尽量大。通过设计这种三元组的形式，可以在样本较少时产生大量的样本三元组对编码器进行训练，从而缓解过拟合问题，进而训练出有效的编码器。

综上，在样本量较小的情况下，通过构建三元组对比学习深度神经网络能够挖掘样本与样本之间的相互关系，提升了编码器的精确度，并结合自

注意力机制考虑不同样本对编码器的贡献程度，进一步挖掘了样本间的关联关系。

步骤 S130，利用经训练的编码器将输入样本数据编码为高维特征，并利用该高维特征训练分类器，其中训练过程中冻结编码器的参数。

在一个实施例中，步骤 S130 包括以下子步骤：

步骤 S131，构建分类器

在获取合适的编码器参数以后，将编码器参数冻结，将输入样本编码为高维特征，再将高维特征输入后续的分类器网络中以获得预测结果，分类器网络如图 4 所示，其整体上包括自注意力模块和多层感知机，多层感知机可采用多层全连接层构建。

具体地，将编码高维特征输入自注意力模块，然后将处理后的特征输入由两层全连接层构成的多层感知机中，其中第一层全连接层的输入神经元数量为 $2048 * N$ ， N 表示输入触觉数据的传感器来源个数，输出神经元数量为 1024，后接一层 Relu 非线性变换激活层；第二层全连接层的输入神经元数量为 1024，输出神经元数量为 2，后接一层 Sigmoid 非线性变换激活层，最终输出分别预测抓取成功或失败的得分。

步骤 S132，训练分类器

使用训练集对上述分类器神经网络用二值交叉熵（Binary Cross-Entropy）损失函数进行训练，二值交叉熵损失函数 L_B 形式化表示如下：

$$L_B = - \left(y \log(p(x)) + (1 - y) \log(1 - p(x)) \right) \quad (2)$$

其中， x 表示输入的样本数据， y 表示该样本实际上抓取成功或失败的类别，例如用 $y=1$ 表示抓取成功， $y=0$ 表示抓取失败， $p(x)$ 表示在当前输入样本数据情况下，分类器预测该样本抓取成功的概率。

在分类器训练完成后，可使用测试集验证该模型的有效性，评价指标为正确预测准确率。

步骤 S140，将经训练的编码器和分类器组合为预测模型，针对实时采集的触觉数据预测机器人的抓取结果。

在完成编码器和分类器的训练后，可将编码器和分类器组合为整体的预测模型，利用该预测模型针对实时采集的触觉数据预测机器臂的抓取是否

成功。

综上所述，相对于现有技术，本发明至少具有以下优势：

1) 本发明提供一种适应任意数量触觉传感器的深度神经网络（包括编码器和分类器等）来通过抓取前的触觉信息预测抓取稳定性，可以在框架基本不变的情况下，结合单个或多个触觉传感器信息场景下进行有效的抓取稳定性预测，且其框架、思路核心思想与触觉传感器、机器人类型无关，具有较好的泛化性。

2) 本发明基于三元组对比学习训练编码器。通过对比学习，以缩小同类样本编码后深度高维特征欧氏距离，扩大异类样本编码后深度高维特征欧氏距离为目标训练特征编码器，以此发掘样本间的关系，并且通过构建三元组产生大量新的训练单元，从而提升了小样本情境下的网络预测准确率表现，从而在样本较少的情况下缓和训练过程中容易产生的过拟合问题，获得较好的编码器参数。通过对比样本，挖掘了样本之间的相互关系，而并不单独考虑样本进行端到端训练。这样的方式使训练过程对样本数据潜在信息的利用更充分，获得更好的训练效果。

3) 本发明在网络中引入自注意力机制。自注意力机制有助于网络在训练过程中挖掘特征内部的相关性，突出高维深度特征中有助于提高分类准确率的重要部分，进一步提升模型的预测准确率。

4) 本发明所提出网络结构已在 Calandra 公开数据集上得到很好的验证，相比现有模型预测性能有所提升。

本发明可以是系统、方法和/或计算机程序产品。计算机程序产品可以包括计算机可读存储介质，其上载有用于使处理器实现本发明的各个方面的计算机可读程序指令。

计算机可读存储介质可以是可以保持和存储由指令执行设备使用的指令的有形设备。计算机可读存储介质例如可以是但不限于电存储设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意合适的组合。计算机可读存储介质的更具体的例子（非穷举的列表）包括：便携式计算机盘、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）、静态随机存取存储器（SRAM）、

便携式压缩盘只读存储器（CD-ROM）、数字多功能盘（DVD）、记忆棒、软盘、机械编码设备、例如其上存储有指令的打孔卡或凹槽内凸起结构、以及上述的任意合适的组合。这里所使用的计算机可读存储介质不被解释为瞬时信号本身，诸如无线电波或者其他自由传播的电磁波、通过波导或其他传输媒介传播的电磁波（例如，通过光纤电缆的光脉冲）、或者通过电线传输的电信号。

这里所描述的计算机可读程序指令可以从计算机可读存储介质下载到各个计算/处理设备，或者通过网络、例如因特网、局域网、广域网和/或无线网下载到外部计算机或外部存储设备。网络可以包括铜传输电缆、光纤传输、无线传输、路由器、防火墙、交换机、网关计算机和/或边缘服务器。每个计算/处理设备中的网络适配卡或者网络接口从网络接收计算机可读程序指令，并转发该计算机可读程序指令，以供存储在各个计算/处理设备中的计算机可读存储介质中。

用于执行本发明操作的计算机程序指令可以是汇编指令、指令集架构（ISA）指令、机器指令、机器相关指令、微代码、固件指令、状态设置数据、或者以一种或多种编程语言的任意组合编写的源代码或目标代码，所述编程语言包括面向对象的编程语言—诸如 Smalltalk、C++、Python 等，以及常规的过程式编程语言—诸如“C”语言或类似的编程语言。计算机可读程序指令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—包括局域网(LAN)或广域网(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）。在一些实施例中，通过利用计算机可读程序指令的状态信息来个性化定制电子电路，例如可编程逻辑电路、现场可编程门阵列（FPGA）或可编程逻辑阵列（PLA），该电子电路可以执行计算机可读程序指令，从而实现本发明的各个方面。

这里参照根据本发明实施例的方法、装置（系统）和计算机程序产品的流程图和/或框图描述了本发明的各个方面。应当理解，流程图和/或框图

的每个方框以及流程图和/或框图中各方框的组合，都可以由计算机可读程序指令实现。

这些计算机可读程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器，从而生产出一种机器，使得这些指令在通过计算机或其它可编程数据处理装置的处理器执行时，产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中，这些指令使得计算机、可编程数据处理装置和/或其他设备以特定方式工作，从而，存储有指令的计算机可读介质则包括一个制造品，其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

也可以把计算机可读程序指令加载到计算机、其它可编程数据处理装置、或其它设备上，使得在计算机、其它可编程数据处理装置或其它设备上执行一系列操作步骤，以产生计算机实现的过程，从而使得在计算机、其它可编程数据处理装置、或其它设备上执行的指令实现流程图和/或框图中的一个或多个方框中规定的功能/动作。

附图中的流程图和框图显示了根据本发明的多个实施例的系统、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、程序段或指令的一部分，所述模块、程序段或指令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。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，两个连续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，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动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。对于本领域技术人员来说公知的是，通过硬件方式实现、通过软件方式实现以及通过软件和硬件结合的方式实现都是等价的。

以上已经描述了本发明的各实施例，上述说明是示例性的，并非穷尽性的，并且也不限于所披露的各实施例。在不偏离所说明的各实施例的范围

和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择，旨在最好地解释各实施例的原理、实际应用或对市场中的技术改进，或者使本技术领域的其它普通技术人员能理解本文披露的各实施例。本发明的范围由所附权利要求来限定。

权 利 要 求 书

1. 一种基于三元组对比网络的机器人抓取预测方法，包括以下步骤：

构建训练集，其中每条样本数据反映机器臂抓取物体过程中多个时刻的触觉数据与分类标签之间的对应关系，该分类标签用于指示抓取是否成功；

基于设定的损失函数训练编码器，获得优化参数，其中在训练过程中，在所述训练集随机抽取两个标签一致的样本分别作为锚样本和正样本，并随机抽取一个标签相反的样本作为负样本；将样本数据分别输入对应编码器中编码得到锚样本的高维特征、正样本的高维特征以及负样本的高维特征；

冻结所述编码器的优化参数，并利用经训练的编码器将输入样本数据编码为高维特征，进而将该高维特征输入分类器进行训练；

利用经训练的编码器和经训练的分类器针对实时采集的触觉数据预测抓取结果。

2. 根据权利要求 1 所述的方法，其特征在于，根据以下步骤构建所述训练集：

采集机器臂抓取物体的触觉数据，包括三个时刻的触觉数据，分别是机器臂置于初始位置时标记为时刻 T_a ，机器臂抓住目标物品但尚未提升时标记为时刻 T_b ，抓住物品在空中悬停设定时间后标记为时刻 T_c ；

通过将 T_a 、 T_b 时刻的触觉数据图像进行随机翻转对数据进行增强，并将触觉数据随机裁剪为设定的尺寸，获得数据集；

从所述数据集中随机选取设定数量的抓取物品种类，并针对每种物品随机选取设定数量的数据，构建为所述训练集。

3. 根据权利要求 1 所述的方法，其特征在于，所述编码器采用三元对比损失函数作为评价指标进行训练，表示为：

$$L_T = \left\| f_a - f_p \right\|_2^2 - \left\| f_a - f_n \right\|_2^2 + 1$$

其中 L_T 表示三元对比损失， f_a 、 f_p 、 f_n 分别表示锚样本、正样本和负样本的高维特征。

4. 根据权利要求 1 所述的方法，其特征在于，利用二值交叉熵损失函数训练所述分类器，表示为：

$$L_B = - \left(y \log(p(x)) + (1 - y) \log(1 - p(x)) \right)$$

其中， x 表示输入的样本数据， y 表示该样本实际上抓取成功或失败的

类别标签， $p(x)$ 表示在当前输入样本数据情况下，预测抓取成功的概率。

5. 根据权利要求 1 所述的方法，其特征在于，所述编码器包括骨干网络，该骨干网络用于从 G_a 与 $G_b - G_a$ 在通道方向叠加的图像中提取高维特征，其中， G_a 、 G_b 分别表示在 T_a 、 T_b 时刻采集到的触觉数据。

6. 根据权利要求 5 所述的方法，其特征在于，所述骨干网络设置为多个，每个骨干网络用于从对应的触觉数据中提取高维特征，并将所提取的高维特征进行拼接后，输入到第一自注意力模块，得到融合高维特征。

7. 根据权利要求 6 所述的方法，其特征在于，所述分类器包括第二自注意力模块和多层感知机，该多层感知机包含两层全连接层，其中第一层全连接层后接一层 Relu 非线性变换激活层；第二层全连接层输出神经元数量为 2，后接一层 Sigmoid 非线性变换激活层，所述分类器输出预测抓取成功或失败的得分。

8. 根据权利要求 5 所述的方法，其特征在于，所述骨干网络基于 Resnet-50、VGG 或 AlexNet 网络构建。

9. 一种计算机可读存储介质，其上存储有计算机程序，其中，该计算机程序被处理器执行时实现根据权利要求 1 至 8 中任一项所述方法的步骤。

10. 一种计算机设备，包括存储器和处理器，在所述存储器上存储有能够在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权利要求 1 至 8 中任一项所述的方法的步骤。

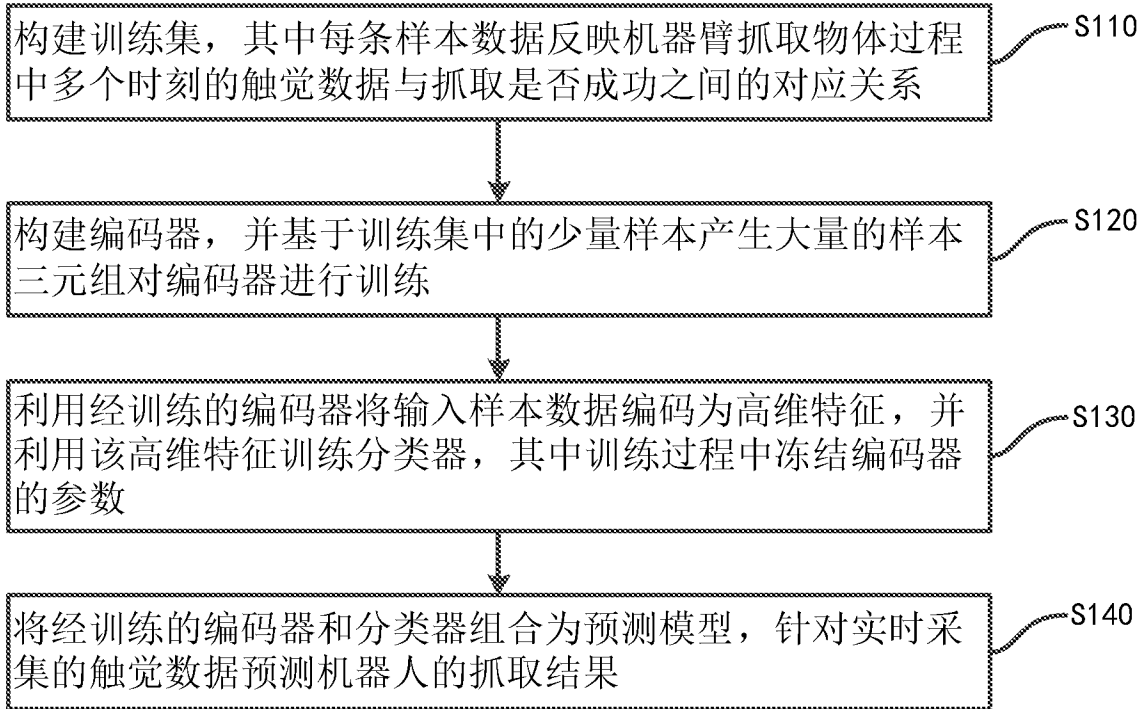


图 1

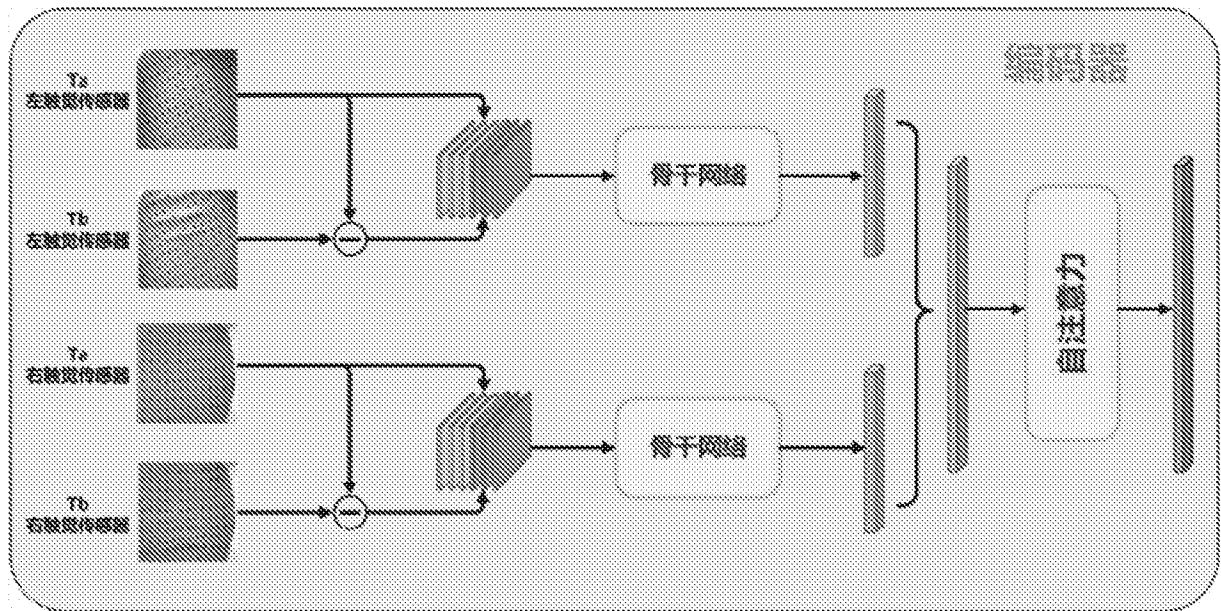


图 2

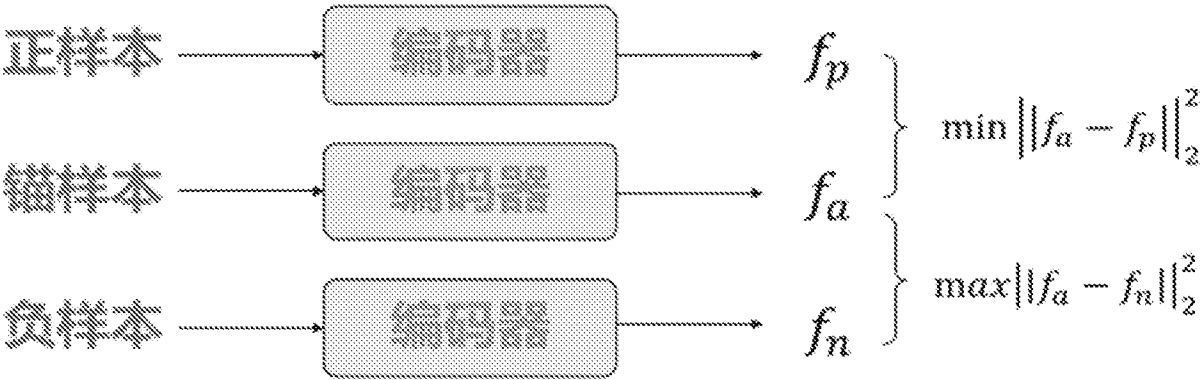


图 3

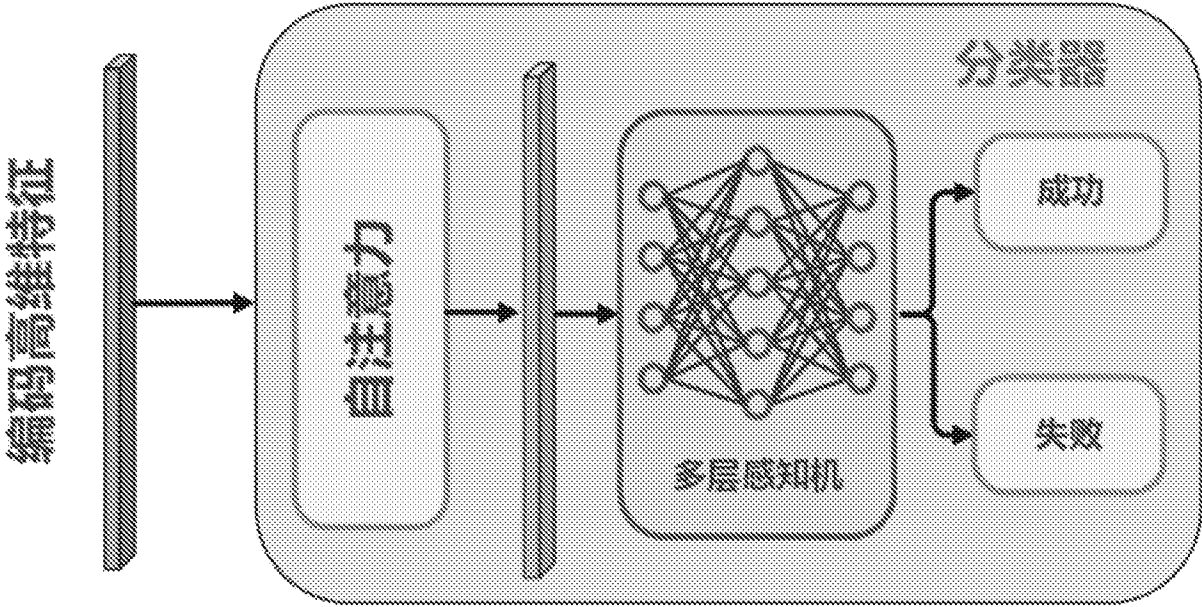


图 4

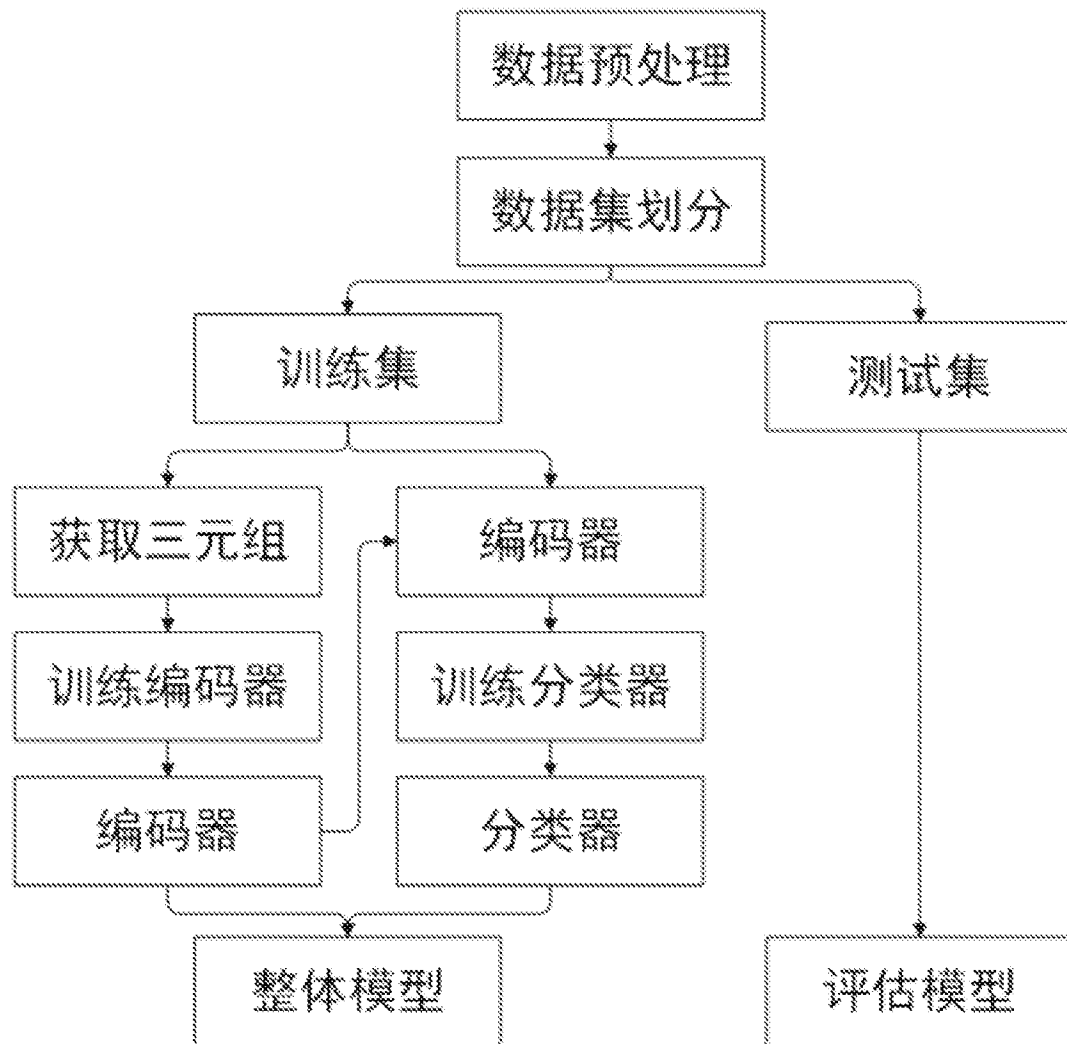


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/137001

A. CLASSIFICATION OF SUBJECT MATTER B25J15/08(2006.01)i; B25J9/16(2006.01)i; B25J13/08(2006.01)i; G06N3/04(2023.01)i; G06V10/80(2022.01)i; G06V10/764(2022.01)i; 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:B25J,G06K,G06N,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, CJFD, WPABSC, VCN, ENTXT, VEN: 深圳先进技术研究院, 机器人, 机械手, 机械臂, 抓取, 预测, 三元, 对比, 神经, 网络, 触觉, 标签, 对应, 编码器, 分类器, 高维, robot, manipulater, ternary, group, grab+, predict+, encoder, classifier																							
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1"> <thead> <tr> <th>Category*</th> <th>Citation of document, with indication, where appropriate, of the relevant passages</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>CN 110691676 A (GOOGLE LLC) 14 January 2020 (2020-01-14) description, paragraphs 0046-0060, and figures 1-6</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 105956351 A (SHANGHAI AEROSPACE CONTROL TECHNOLOGY INSTITUTE) 21 September 2016 (2016-09-21) entire document</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 112733965 A (XI'AN UNIVERSITY OF TECHNOLOGY) 30 April 2021 (2021-04-30) entire document</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 105005787 A (TSINGHUA UNIVERSITY) 28 October 2015 (2015-10-28) entire document</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 113610151 A (HARBIN UNIVERSITY OF SCIENCE AND TECHNOLOGY) 05 November 2021 (2021-11-05) entire document</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 112668607 A (SHENZHEN INSTITUTE OF ADVANCED TECHNOLOGY) 16 April 2021 (2021-04-16) entire document</td> <td>1-10</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	A	CN 110691676 A (GOOGLE LLC) 14 January 2020 (2020-01-14) description, paragraphs 0046-0060, and figures 1-6	1-10	A	CN 105956351 A (SHANGHAI AEROSPACE CONTROL TECHNOLOGY INSTITUTE) 21 September 2016 (2016-09-21) entire document	1-10	A	CN 112733965 A (XI'AN UNIVERSITY OF TECHNOLOGY) 30 April 2021 (2021-04-30) entire document	1-10	A	CN 105005787 A (TSINGHUA UNIVERSITY) 28 October 2015 (2015-10-28) entire document	1-10	A	CN 113610151 A (HARBIN UNIVERSITY OF SCIENCE AND TECHNOLOGY) 05 November 2021 (2021-11-05) entire document	1-10	A	CN 112668607 A (SHENZHEN INSTITUTE OF ADVANCED TECHNOLOGY) 16 April 2021 (2021-04-16) entire document	1-10
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.																					
A	CN 110691676 A (GOOGLE LLC) 14 January 2020 (2020-01-14) description, paragraphs 0046-0060, and figures 1-6	1-10																					
A	CN 105956351 A (SHANGHAI AEROSPACE CONTROL TECHNOLOGY INSTITUTE) 21 September 2016 (2016-09-21) entire document	1-10																					
A	CN 112733965 A (XI'AN UNIVERSITY OF TECHNOLOGY) 30 April 2021 (2021-04-30) entire document	1-10																					
A	CN 105005787 A (TSINGHUA UNIVERSITY) 28 October 2015 (2015-10-28) entire document	1-10																					
A	CN 113610151 A (HARBIN UNIVERSITY OF SCIENCE AND TECHNOLOGY) 05 November 2021 (2021-11-05) entire document	1-10																					
A	CN 112668607 A (SHENZHEN INSTITUTE OF ADVANCED TECHNOLOGY) 16 April 2021 (2021-04-16) 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 07 June 2023		Date of mailing of the international search report 12 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

International application No.

PCT/CN2022/137001

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2019196436 A1 (X DEVELOPMENT LLC) 27 June 2019 (2019-06-27) entire document	1-10
A	CN 111459278 A (AEROSPACE INFORMATION RESEARCH INSTITUTE, CHINESE ACADEMY OF SCIENCES) 28 July 2020 (2020-07-28) entire document	1-10
A	CN 113327295 A (SOUTH CHINA UNIVERSITY OF TECHNOLOGY) 31 August 2021 (2021-08-31) entire document	1-10
A	CN 113168567 A (HRL LABORATORIES, LLC) 23 July 2021 (2021-07-23) entire document	1-10
A	CN 115034286 A (NATIONAL COMPUTER NETWORK AND INFORMATION SECURITY MANAGEMENT CENTER) 09 September 2022 (2022-09-09) entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/137001

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110691676	A	14 January 2020	EP	3515671	A1	31 July 2019
				EP	3515671	B1	13 May 2020
				US	2021053217	A1	25 February 2021
				US	11554483	B2	17 January 2023
				EP	3693138	A1	12 August 2020
				EP	3693138	B1	03 August 2022
				WO	2018236753	A1	27 December 2018
				US	2020094405	A1	26 March 2020
				US	10864631	B2	15 December 2020
CN	105956351	A	21 September 2016	None			
CN	112733965	A	30 April 2021	None			
CN	105005787	A	28 October 2015	None			
CN	113610151	A	05 November 2021	None			
CN	112668607	A	16 April 2021	None			
US	2019196436	A1	27 June 2019	WO	2019126323	A1	27 June 2019
				US	2020348642	A1	05 November 2020
				US	10955811	B2	23 March 2021
				US	10754318	B2	25 August 2020
CN	111459278	A	28 July 2020	None			
CN	113327295	A	31 August 2021	None			
CN	113168567	A	23 July 2021	EP	3874417	A1	08 September 2021
				US	2020130177	A1	30 April 2020
				WO	2020091871	A1	07 May 2020
CN	115034286	A	09 September 2022	None			

国际检索报告

国际申请号

PCT/CN2022/137001

A. 主题的分类

B25J15/08(2006.01)i; B25J9/16(2006.01)i; B25J13/08(2006.01)i; G06N3/04(2023.01)i; G06V10/80(2022.01)i; G06V10/764(2022.01)i; 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:B25J, G06K, G06N, 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))

CNXT, CJFD, WPABSC, VCN, ENTXT, VEN; 深圳先进技术研究院, 机器人, 机械手, 机械臂, 抓取, 预测, 三元, 对比, 神经, 网络, 触觉, 标签, 对应, 编码器, 分类器, 高维, robot, manipulat+, ternary, group, grab+, predict+, encoder, classifier

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 110691676 A (谷歌有限责任公司) 2020年1月14日 (2020 - 01 - 14) 说明书第0046-0060段, 图1-6	1-10
A	CN 105956351 A (上海航天控制技术研究所) 2016年9月21日 (2016 - 09 - 21) 全文	1-10
A	CN 112733965 A (西安理工大学) 2021年4月30日 (2021 - 04 - 30) 全文	1-10
A	CN 105005787 A (清华大学) 2015年10月28日 (2015 - 10 - 28) 全文	1-10
A	CN 113610151 A (哈尔滨理工大学) 2021年11月5日 (2021 - 11 - 05) 全文	1-10
A	CN 112668607 A (深圳先进技术研究院) 2021年4月16日 (2021 - 04 - 16) 全文	1-10
A	US 2019196436 A1 (X DEV. L.L.C.) 2019年6月27日 (2019 - 06 - 27) 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

2023年6月7日

国际检索报告邮寄日期

2023年6月12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

马琳

电话号码 (+86) 010-53961015

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 111459278 A (中国科学院空天信息创新研究院) 2020年7月28日 (2020 - 07 - 28) 全文	1-10
A	CN 113327295 A (华南理工大学) 2021年8月31日 (2021 - 08 - 31) 全文	1-10
A	CN 113168567 A (HRL实验室有限责任公司) 2021年7月23日 (2021 - 07 - 23) 全文	1-10
A	CN 115034286 A (国家计算机网络与信息安全管理中心) 2022年9月9日 (2022 - 09 - 09) 全文	1-10

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/137001

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	110691676	A	2020年1月14日	EP	3515671	A1	2019年7月31日
				EP	3515671	B1	2020年5月13日
				US	2021053217	A1	2021年2月25日
				US	11554483	B2	2023年1月17日
				EP	3693138	A1	2020年8月12日
				EP	3693138	B1	2022年8月3日
				WO	2018236753	A1	2018年12月27日
				US	2020094405	A1	2020年3月26日
				US	10864631	B2	2020年12月15日
CN	105956351	A	2016年9月21日	无			
CN	112733965	A	2021年4月30日	无			
CN	105005787	A	2015年10月28日	无			
CN	113610151	A	2021年11月5日	无			
CN	112668607	A	2021年4月16日	无			
US	2019196436	A1	2019年6月27日	WO	2019126323	A1	2019年6月27日
				US	2020348642	A1	2020年11月5日
				US	10955811	B2	2021年3月23日
				US	10754318	B2	2020年8月25日
CN	111459278	A	2020年7月28日	无			
CN	113327295	A	2021年8月31日	无			
CN	113168567	A	2021年7月23日	EP	3874417	A1	2021年9月8日
				US	2020130177	A1	2020年4月30日
				WO	2020091871	A1	2020年5月7日
CN	115034286	A	2022年9月9日	无			