

(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
2022 年 2 月 24 日 (24.02.2022)



(10) 国际公布号
WO 2022/037233 A1

(51) 国际专利分类号:
G06K 9/00 (2006.01)

(21) 国际申请号: PCT/CN2021/101128

(22) 国际申请日: 2021 年 6 月 21 日 (21.06.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010830474.4 2020 年 8 月 18 日 (18.08.2020) CN

(71) 申请人: 浙江大学(ZHEJIANG UNIVERSITY) [CN/
CN]; 中国浙江省杭州市西湖区余杭塘路
866 号, Zhejiang 310058 (CN)。

(72) 发明人: 宋杰(SONG, Jie); 中国浙江省杭州市西湖区余杭塘路 866 号, Zhejiang 310058 (CN)。 宋明黎(SONG, Mingli); 中国浙江省杭州市西湖区余杭塘路 866 号, Zhejiang 310058 (CN)。 冯尊磊(FENG, Zunlei); 中国浙江省杭州市西湖区余杭塘路 866 号, Zhejiang 310058 (CN)。 陈刚(CHEN, Gang); 中国浙江省杭州市西湖区余杭塘路 866 号, Zhejiang 310058 (CN)。

(74) 代理人: 杭州求是专利事务所有限公司(HANGZHOU QIUSHI PATENT OFFICE CO., LTD.); 中国浙江省杭州市西湖区古墩路 701 号绿城紫金广场 C 座 1506 室邱启旺, Zhejiang 310014 (CN)。

(54) Title: SMALL SAMPLE VISUAL TARGET IDENTIFICATION METHOD BASED ON SELF-SUPERVISED KNOWLEDGE TRANSFER

(54) 发明名称: 一种基于自监督知识迁移的小样本视觉目标识别方法

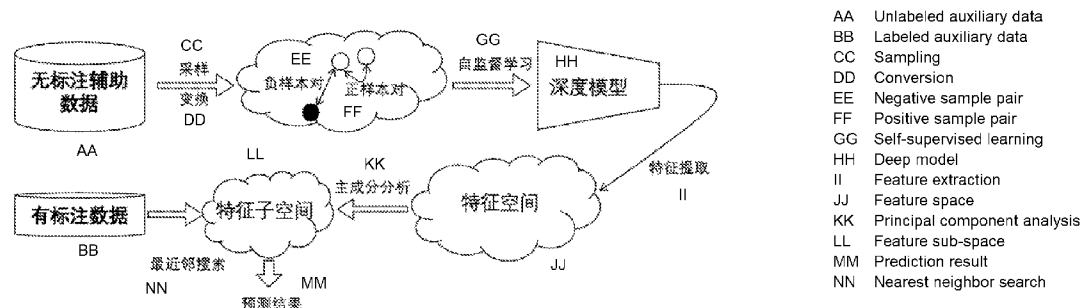


图 1

(57) Abstract: A small sample visual target identification method based on self-supervised knowledge transfer. The method comprises: 1) collecting unlabeled auxiliary data which is weakly related to a task and has a relatively large amount of data, and labeled target data which is strongly related to the task and has a relatively small amount of data; 2) constructing, on the unlabeled auxiliary data with a relatively large amount of data and by means of data conversion, a positive sample pair and a negative sample pair, and performing self-supervised learning by using a contrast loss function, so as to pre-train a deep neural network; 3) extracting a feature of target data by using a pre-trained model, and performing data dimension reduction on the basis of a feature space, so as to learn a feature of a feature sub-space with a relatively strong discriminative ability regarding the target data; and 4) using a feature expression, in the sub-space, of a small amount of labeled data of each category as a feature prototype of the category, and performing classification prediction on test data by using a nearest neighbor method.

(57) 摘要: 一种基于自监督知识迁移的小样本视觉目标识别方法, 包括: 1) 收集与任务弱相关、数量较多的无标注辅助数据, 以及与任务强相关、数据量较小的有标注目标数据; 2) 在数据量较多的无标注辅助数据上通过数据变换构造正样本对以及负样本对, 并利用对比损失函数作自监督学习, 预训练一个深度神经网络; 3) 利用预训练模型提取目标数据的特征, 并在该特征空间的基础进行数据降维, 学习到一个对目标数据具有较强判别能力的特征子空间; 4) 将每个类别少量标注数据在该子空间的特征表达作为该类别的特征原型, 利用最近邻方法对测试数据进行分类预测。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

一种基于自监督知识迁移的小样本视觉目标识别方法

技术领域

本发明属于计算机视觉领域，针对图像分类、目标检测、视频监控等场景中标注数据量不足的问题，提出了一种基于自监督知识迁移的小样本视觉目标识别方法。

背景技术

计算机视觉目标识别是计算机视觉以及人工智能领域最基础、最核心的问题之一。近几年由于深度学习技术的发展，视觉目标识别已经取得了非常巨大的进步。现有的深度模型在海量标注数据驱动下，已经能够实现单个模型准确识别上千种视觉目标。然而在视觉目标分类、检测等场景中，现有深度模型依赖大量的高质量标注数据进行学习。但现实生活场景中由于数据的采集成本以及标注成本普遍较高，通常很难获取到大规模标注数据用以模型训练。在数据较少情况下，这些深度模型通常很难进行有效学习。

现有的解决小样本问题的方法主要通过元学习来实现：（1）首先在大量相关视觉分类任务上进行有监督学习来提取任务通用的信息；（2）其次利用这些通用信息引导模型对当前的视觉分类任务进行学习。元学习依然需要大量有标注的相关视觉数据进行学习，数据标注成本依然较高；其次元学习对辅助数据要求较高，需要辅助数据与目标数据强相关，否则学习到的知识很难引导模型对目标数据的有效学习。

发明内容

针对现实场景中视觉分类任务中数据量不足的现状，本发明提出了一种基于自监督知识迁移的小样本视觉目标识别方法，从无标注、弱相关的辅助数据学习通用知识解决目标分类任务。

为实现上述目的，本发明所述的基于自监督知识迁移的小样本视觉目标识别方法，包括如下步骤：

1) 收集与任务弱相关、数量较多的无标注辅助数据，以及与任务强相关、数据量较小的有标注目标数据；

2) 在数据量较多的无标注辅助数据上通过数据变换构造正样本对以及负样本对，并利用对比损失函数作自监督学习，预训练一个深度模型；

3) 利用预训练模型提取目标数据的特征，并在该特征空间的基础进行数据降维，学习到一个对目标数据具有较强判别能力的特征子空间；

4) 将每个类别少量标注数据在该子空间的特征表达作为该类别的特征原型，利用最近邻方法对测试数据进行分类预测。

进一步，步骤 1) 所述的辅助数据无任何标注，且数据量充足能够用以训练；所述的目标数据有标注，但数据量较少，通常每个类仅有几张标注图像。

进一步，步骤 2) 所述的深度模型的最终目标是对目标数据进行识别，即输出一个目标数据图像中视觉物体的类别。

进一步，步骤 1) 所述的辅助数据用以深度模型对通用知识的学

习；所述的目标数据用以引导深度模型对特定目标数据的快速迁移和准确识别。

进一步，步骤 2) 所述的自监督学习指的是模型的学习无需任何人工标注，仅依赖利用数据本身的自洽信息构造伪任务，以帮助深度模型进行预学习。

进一步，步骤 2) 所述的对比损失通过下式计算：

$$\mathcal{L}_{i,j} = -\log \frac{\exp(\text{sim}(z_i, z_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{I}_{[k \neq i]} \exp(\text{sim}(z_i, z_k)/\tau)} \quad (1)$$

其中， i, j 表示一个正样本对， N 代表模型训练过程中一次迭代更新处理的批大小， z 代表数据的特征向量， τ 代表温度超参数，函数 sim 代表特征相似性： $\text{sim}(u, v) = u^T v / \|u\| \|v\|$ 。

进一步，步骤 3) 所述的对于特征的数据降维通过主成分分析 PCA、ICA 等手段进行。

进一步，步骤 4) 所述的每个类别在特征子空间内的特征原型为该类别数据的特征均值，其计算公式为

$$z = \frac{1}{K} \sum_{k=1}^K z_k \quad (2)$$

其中， K 为有标注数据中每个类别的标注数据量。

本发明的有益效果如下：

本发明是一种基于自监督知识迁移的小样本视觉目标识别方法。它利用无标注、弱相关的辅助数据进行自学习，然后利用有标注、小样本的目标数据进行特征降维以及目标识别。

与传统方法相比，本发明能够有效缓解深度模型对有标注训练数据的依赖性，提高算法在现实场景下实用性。

附图说明

图 1 是本发明方法的流程示意图。

具体实施方式

下面结合附图和具体实施例对本发明的技术方案进行清晰、完整的解释和描述。

生物物种识别任务中，由于数据的采集成本以及标注成本普遍较高，通常很难获取到大规模标注数据用以模型训练。在数据较少情况下，现有深度模型通常很难进行有效学习。

针对生物物种识别任务中数据量不足的现状，本发明提出了一种基于自监督知识迁移的小样本生物物种识别方法，从无标注、弱相关的辅助数据学习通用知识解决物种分类任务，具体包括如下步骤：

1) 收集与待识别物种弱相关、且无需标注的辅助图像数据，以及少量有标注的待识别物种图像数据；

2) 在数据量较多的无标注辅助图像数据上通过数据变换构造正样本对以及负样本对，并利用对比损失函数作自监督学习，预训练一个深度模型；

3) 利用预训练模型提取待识别物种的特征，并在该特征空间的基础进行数据降维，学习到一个对物种识别具有较强判别能力的特征子空间；

4) 将每个物种少量标注数据在该子空间的特征表达作为该类别的特征原型，利用最近邻搜索方法对测试数据进行预测。

步骤 1) 所述的辅助数据无任何标注，且数据量充足能够用以训

练；所述的生物物种数据有标注，但数据量较少，通常每个物种仅有几张标注图像。

步骤2)所述的深度模型的最终目标是对生物物种进行识别，即输出一个图像中生物物种的名称。

步骤1)所述的辅助数据用以深度模型对通用知识的学习；所述的目标数据用以引导深度模型对特定目标数据的快速迁移和准确识别。

步骤2)所述的自监督学习指的是模型的学习无需任何人工标注，仅依赖利用数据本身的自洽信息构造伪任务，以帮助深度模型进行预学习。

步骤2)所述的对比损失通过下式计算：

$$\mathcal{L}_{i,j} = -\log \frac{\exp(\text{sim}(z_i, z_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{I}_{[k \neq i]} \exp(\text{sim}(z_i, z_k)/\tau)} \quad (1)$$

其中， i, j 表示一个正样本对， N 代表模型训练过程中一次迭代更新处理的批大小， z 代表数据的特征向量， τ 代表温度超参数，函数 sim 代表特征相似性： $\text{sim}(u, v) = u^T v / \|u\| \|v\|$ 。

步骤3)所述的对于特征的数据降维通过主成分分析 PCA、ICA等手段进行。

步骤4)所述的每个物种在特征子空间内的特征原型为该物种数据的特征均值，其计算公式为

$$z = \frac{1}{K} \sum_{k=1}^K z_k \quad (2)$$

其中， K 为有标注数据中每个物种的标注数据量。

具体地，本发明提出了一种基于自监督知识迁移的小样本生物物

种识别方法，该方法能够从利用自监督学习通用的知识，并利用该知识解决小样本条件下生物物种识别问题。包括如下步骤：

步骤 1，采集无标注辅助图像数据和有标注的生物物种数据。其中无标注辅助数据 100 万张图片；生物物种数据中每个类别包含 $K = 5$ 张图片。所有图像块放缩至 224×224 。

步骤 2，构建正样本对和负样本对，并利用自监督学习训练深度模型。每次训练迭代中，从辅助数据中随机选取 $N = 32$ 张图片，并通过数据转换函数构造另外 32 个数据，构成大小为 64 的批数据。正样本对由对应的数据对构成，其它数据对构成负样本对。温度超参数为 τ 设置为 4，

步骤 3，利用预训练模型进行特征提取，并利用生物物种数据进行数据降维。模型预训练后，利用该模型提取所有生物物种数据的特征，并基于该特征利用主成分分析对特征空间进行降维，只保留前 5% 主成分。

步骤 4，对生物物种图像进行分类。在特征子空间中计算每个物种的特征均值，并利用该特征均值对每个图像的特征做最近邻搜索，选择距离最近的那个物种作为该测试图像的预测结果。

本说明书实施例所述的内容仅仅是对发明构思的实现形式的列举，本发明的保护范围的不应当被视为仅限于实施例所陈述的具体形式，本发明的保护范围也及于本领域技术人员根据本发明构思所能够想到的等同技术手段。

权 利 要 求 书

1. 一种基于自监督知识迁移的小样本视觉目标识别方法，包含如下步骤：

1) 收集与任务弱相关、数量较多的无标注辅助数据，以及与任务强相关、数据量较小的有标注目标数据；

2) 在数据量较多的无标注辅助数据上通过数据变换构造正样本对以及负样本对，并利用对比损失函数作自监督学习，预训练一个深度模型；

3) 利用预训练模型提取目标数据的特征，并在该特征空间的基础进行数据降维，学习到一个对目标数据具有较强判别能力的特征子空间；

4) 将每个类别少量标注数据在该子空间的特征表达作为该类别的特征原型，利用最近邻方法对测试数据进行分类预测。

2. 根据权利要求 1 所述的一种基于自监督知识迁移的小样本视觉目标识别方法，其特征在于：步骤 1) 所述的辅助数据无任何标注，且数据量充足能够用以训练；步骤 1) 所述的目标数据有标注，但数据量较少，通常每个类仅有几张标注图像。

3. 根据权利要求 1 所述的一种基于自监督知识迁移的小样本视觉目标识别方法，其特征在于：步骤 2) 所述的深度模型的最终目标是对目标数据进行识别，即输出一个目标数据图像中视觉物体的类别。

4. 根据权利要求 1 所述的一种基于自监督知识迁移的小样本视觉目标识别方法，其特征在于：步骤 1) 所述的辅助数据用以深度模型对通用知识的学习，该部分知识能够泛化到多个任务中；所述的目

标数据用以引导深度模型对特定目标数据的快速迁移和准确识别。

5. 根据权利要求 1 所述的一种基于自监督知识迁移的小样本视觉目标识别方法，其特征在于：步骤 2) 所述的自监督学习指的是模型的学习无需任何人工标注，仅依赖利用数据本身的自洽信息构造伪任务，以帮助深度模型进行预学习。

6. 根据权利要求 1 所述的一种基于自监督知识迁移的小样本视觉目标识别方法，其特征在于：步骤 2) 所述的对比损失函数通过下式计算：

$$\mathcal{L}_{i,j} = -\log \frac{\exp(\text{sim}(z_i, z_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{I}_{[k \neq i]} \exp(\text{sim}(z_i, z_k)/\tau)} \quad (1)$$

其中， i, j 表示一个正样本对， N 代表模型训练过程中一次迭代更新处理的批大小， z 代表数据的特征向量， τ 代表温度超参数，函数 sim 代表特征相似性： $\text{sim}(u, v) = u^T v / \|u\| \|v\|$ 。

7. 根据权利要求 1 所述的一种基于自监督知识迁移的小样本视觉目标识别方法，其特征在于：步骤 3) 所述的对于特征的数据降维通过主成分分析 PCA、ICA 手段进行。

8. 根据权利要求 1 所述的一种基于自监督知识迁移的小样本视觉目标识别方法，其特征在于：步骤 4) 所述的每个类别在特征子空间内的特征原型为该类别数据的特征均值，其计算公式为

$$z = \frac{1}{K} \sum_{k=1}^K z_k \quad (2)$$

其中， K 为有标注数据中每个类别的标注数据量。

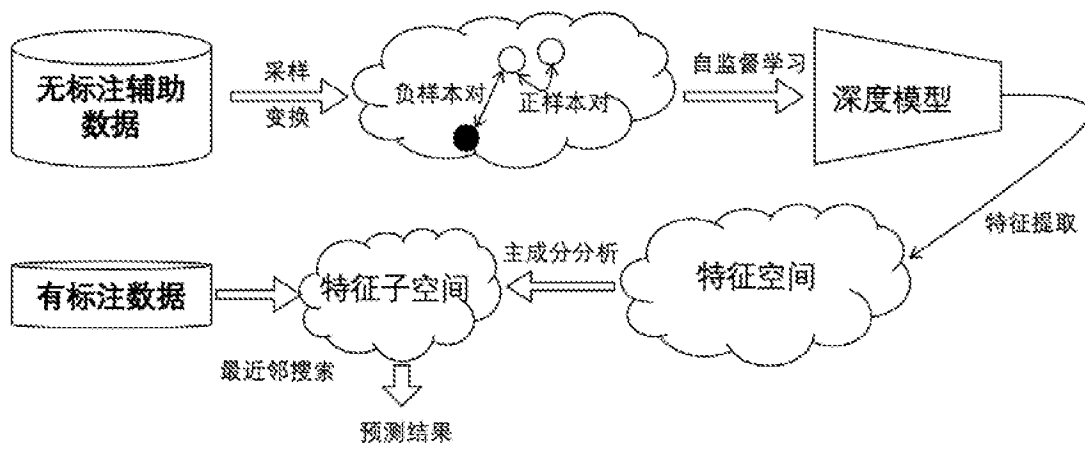


图 1

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/101128

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

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

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

CNABS, CNTXT, VEN, USTXT, WOTXT, EPTXT, CNKI: 自监督, 知识迁移, 无标注, 有标注, 数据, 学习, 神经网络, 深度, 特征, 分类, 训练, 样本, self-supervised, knowledge transfer, non-marked, mark, data, learn, neural network, NN, depth, characteristic, classify, train, 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 112069921 A (ZHEJIANG UNIVERSITY) 11 December 2020 (2020-12-11) description paragraphs 4-48, figure 1	1-8
A	CN 110363231 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 22 October 2019 (2019-10-22) entire document	1-8
X	US 2020160177 A1 (ROYAL BANK OF CANADA) 21 May 2020 (2020-05-21) entire document main body, chapters 2-4	1-8



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

“&” document member of the same patent family

Date of the actual completion of the international search

12 September 2021

Date of mailing of the international search report

24 September 2021

Name and mailing address of the ISA/CN

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/101128

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	112069921	A	11 December 2020	None			
CN	110363231	A	22 October 2019	None			
US	2020160177	A1	21 May 2020	CA	3061717	A1	16 May 2020

国际检索报告

国际申请号

PCT/CN2021/101128

A. 主题的分类

G06K 9/00 (2006.01) i

按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

G06K

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))

CNABS, CNTXT, VEN, USTXT, WOTXT, EPTXT, CNKI: 自监督, 知识迁移, 无标注, 有标注, 数据, 学习, 神经网络, 深度, 特征, 分类, 训练, 样本, self-supervised, knowledge transfer, non-marked, mark, data, learn, neural network, NN, depth, characteristic, classify, train, 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 112069921 A (浙江大学) 2020年 12月 11日 (2020 - 12 - 11) 说明书第4-48段, 图1	1-8
A	CN 110363231 A (平安科技深圳有限公司) 2019年 10月 22日 (2019 - 10 - 22) 全文	1-8
X	US 2020160177 A1 (ROYAL BANK OF CANADA) 2020年 5月 21日 (2020 - 05 - 21) 全文 正文第2-4章	1-8

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2021年 9月 12日

国际检索报告邮寄日期

2021年 9月 24日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)

中国 北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

授权官员

古志春

电话号码 86-(20)-28950386

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/101128

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	112069921	A	2020年 12月 11日	无			
CN	110363231	A	2019年 10月 22日	无			
US	2020160177	A1	2020年 5月 21日	CA	3061717	A1	2020年 5月 16日