

(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



(10) 国际公布号

WO 2020/098121 A1

(43) 国际公布日

2020 年 5 月 22 日 (22.05.2020)

(51) 国际专利分类号:

G06K 9/62 (2006.01)

(21) 国际申请号:

PCT/CN2018/125592

(22) 国际申请日:

2018 年 12 月 29 日 (29.12.2018)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

201811348231.6 2018 年 11 月 13 日 (13.11.2018) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];

中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

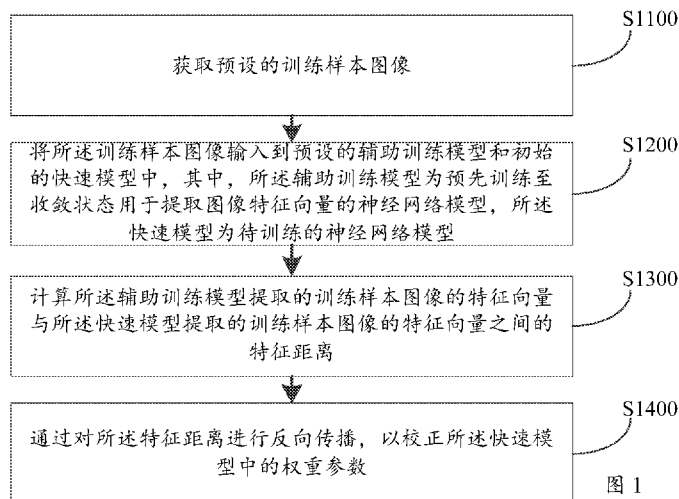
(72) 发明人: 徐玲玲(XU, Lingling); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市立智方成知识产权代理事务所(普通合伙)(LIFANG & PARTNERS LTD.); 中国广东省深圳市福田区金田路 4018 号安联大厦 22D03, Guangdong 518000 (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, 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, 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,

(54) Title: METHOD AND DEVICE FOR TRAINING FAST MODEL, COMPUTER APPARATUS, AND STORAGE MEDIUM

(54) 发明名称: 快速模型的训练方法、装置、计算机设备及存储介质



S1100 Acquire a pre-configured training sample image
S1200 Input the training sample image into a pre-configured auxiliary training model and an initial fast model, wherein the auxiliary training model is a pre-trained neural network model in a convergence state and used to extract an image feature vector, the fast model being a neural network model to be trained
S1300 Calculate a feature distance between a feature vector of the training sample image extracted by the auxiliary training model and a feature vector of the training sample image extracted by the fast model
S1400 Perform back propagation on the feature distance to correct a weight parameter in the fast model

(57) Abstract: Embodiments of the present application disclose a method and device for training a fast model, a computer apparatus, and a storage medium. The method comprises: acquiring a pre-configured training sample image; inputting the training sample image into a pre-configured auxiliary training model and an initial fast model; calculating a feature distance between a feature vector of the training sample image extracted by the auxiliary training model and a feature vector of the training sample image extracted by the fast model; and performing back propagation on the feature distance to correct a weight parameter in the fast model. When the fast model is

PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, 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)。

根据细则4.17的声明:

— 发明人资格 (细则4.17(iv))

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

trained, a sample image involved in training does not need to be marked, thereby reducing the time and effort required for marking, and increasing training speed. A distance between a feature vector representing a sample image feature and output by the auxiliary model and a feature vector representing a sample image feature and output by the fast model is directly calculated, and back propagation is performed, thereby maximally shortening the training time.

(57) 摘要: 本申请实施例公开了一种快速模型的训练方法、装置、计算机设备及存储介质, 包括: 获取预设的训练样本图像; 将所述训练样本图像输入到预设的辅助训练模型和初始的快速模型中; 计算所述辅助训练模型提取的训练样本图像的特征向量与所述快速模型提取的训练样本图像的特征向量之间的特征距离; 通过对所述特征距离进行反向传播, 以校正所述快速模型中的权重参数。通过在对快速模型进行训练时, 无需对参与训练的样本图像进行标记, 节约了标记所需要的时间和花费的精力, 提高了训练的速度。直接计算辅助模型输出的表征样本图像特征的特征向量, 与快速模型输出的表征样本图像特征的特征向量之间的距离并进行反向传播, 能够最大限度的缩短训练时间。

快速模型的训练方法、装置、计算机设备及存储介质

本申请要求于 2018 年 11 月 13 日提交中国专利局、申请号为 201811348231.6，发明名称为“快速模型的训练方法、装置、计算机设备及存储介质”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。

技术领域

本申请实施例涉及模型训练领域，尤其是一种快速模型的训练方法、装置、计算机设备及存储介质。

背景技术

模拟人类实际神经网络的数学方法问世以来，人们已慢慢习惯了把这种人工神经网络直接称为神经网络。神经网络在系统辨识、模式识别、智能控制等领域有着广泛而吸引人的前景，特别在智能控制中，人们对神经网络的自学习功能尤其感兴趣，并且把神经网络这一重要特点看作是解决自动控制中控制器适应能力这个难题的关键钥匙之一。

现有技术中，为了使神经网络模型具有对某一类或者多个类型的图像具有准确分类的能力，需要对初始化的神经网络模型进行训练。其中，训练的方式为：收集一定量级的样本图像，为了使训练至收敛的神经网络模型的鲁棒性更好，需要的训练样本往往是海量的。训练时首先对训练样本进行人为标定，然后将样本图像输入到神经网络模型中，得到神经网络模型输出的分类结果。比对分类结果和人为标定是否一致，不一致时通过反向算法校正神经网络模型的权值，使神经网络模型逐渐的收敛，由于，模型训练时输出的随机性很强，因此，训练过程极为漫长。

本申请的发明人在研究中发现，现有技术中为提高神经网络模型的准确性，往往需要投入大量的样本图像花费大量的时间才能够将神经网络模型训练至收敛状态，而在不具有上述条件的环境下，神经网络模型往往无法进行训练，或者训练得到的神经网络模型稳定性差准确性较低。

发明内容

本申请实施例提供能够通过辅助训练模型在少量样本和较短时间能够将模型，训练至收敛的快速模型的训练方法、装置、计算机设备及存储介质。

为解决上述技术问题，本申请创造的实施例采用的一个技术方案是：提供一种快

速模型的训练方法，包括：获取预设的训练样本图像；将所述训练样本图像输入到预设的辅助训练模型和初始的快速模型中，其中，所述辅助训练模型为预先训练至收敛状态用于提取图像特征向量的神经网络模型，所述快速模型为待训练的神经网络模型；计算所述辅助训练模型提取的训练样本图像的特征向量与所述快速模型提取的训练样本图像的特征向量之间的特征距离；通过对所述特征距离进行反向传播，以校正所述快速模型中的权重参数。

为解决上述技术问题，本申请实施例还提供一种快速模型的训练装置，包括：获取模块，用于获取预设的训练样本图像；处理模块，用于将所述训练样本图像输入到预设的辅助训练模型和初始的快速模型中，其中，所述辅助训练模型为预先训练至收敛状态用于提取图像特征向量的神经网络模型，所述快速模型为待训练的神经网络模型；计算模块，用于计算所述辅助训练模型提取的训练样本图像的特征向量与所述快速模型提取的训练样本图像的特征向量之间的特征距离；执行模块，用于通过对所述特征距离进行反向传播，以校正所述快速模型中的权重参数。

为解决上述技术问题，本申请实施例还提供一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行上述所述快速模型的训练方法的步骤。

为解决上述技术问题，本申请实施例还提供一种存储有计算机可读指令的存储介质，所述计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行上述所述快速模型的训练方法的步骤。

通过在对快速模型进行训练时，无需对参与训练的样本图像进行标记，节约了标记所需要的时间和花费的精力，提高了训练的速度。同时，直接计算辅助模型输出的表征样本图像特征的特征向量，与快速模型输出的表征样本图像特征的特征向量之间的距离（欧式距离和/或余弦距离）并进行反向传播，采用这种方法将快速模型的训练转化为一种单纯的回归算法，能够最大限度的缩短训练时间，且能够保证训练完成时快速模型输出的准确率。

附图说明

图1为本申请实施例快速模型的训练方法的基本流程示意图；

图2为本申请实施例通过阈值确定是否进行反向传播的流程示意图；

图3为本申请实施例计算两个模型提取特征向量之间的特征距离的一种流程示意图；

图4为本申请实施例根据具体应用场景获取训练样本图像的流程示意图；

图 5 为本申请实施例数据库检索所获取训练样本图像的一种流程示意图；

图 6 为本申请实施例派生样本图像的生成方法流程示意图；

图 7 为本申请实施例快速模型的训练装置基本结构示意图；

图 8 为本申请实施例计算机设备基本结构框图。

具体实施方式

为了使本技术领域的人员更好地理解本申请方案，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。

在本申请的说明书和权利要求书及上述附图中的描述的一些流程中，包含了按照特定顺序出现的多个操作，但是应该清楚了解，这些操作可以不按照其在本文中出现的顺序来执行或并行执行，操作的序号如 101、102 等，仅仅是用于区分开各个不同的操作，序号本身不代表任何的执行顺序。另外，这些流程可以包括更多或更少的操作，并且这些操作可以按顺序执行或并行执行。需要说明的是，本文中的“第一”、“第二”等描述，是用于区分不同的消息、设备、模块等，不代表先后顺序，也不限定“第一”和“第二”是不同的类型。

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，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

本技术领域技术人员可以理解，这里所使用的“终端”、“终端设备”既包括无线信号接收器的设备，其仅具备无发射能力的无线信号接收器的设备，又包括接收和发射硬件的设备，其具有能够在双向通信链路上，执行双向通信的接收和发射硬件的设备。这种设备可以包括：蜂窝或其他通信设备，其具有单线路显示器或多线路显示器或没有多线路显示器的蜂窝或其他通信设备；PCS（Personal Communications Service，个人通信系统），其可以组合语音、数据处理、传真和/或数据通信能力；PDA（Personal Digital Assistant，个人数字助理），其可以包括射频接收器、寻呼机、互联网/内联网访问、网络浏览器、记事本、日历和/或 GPS（Global Positioning System，全球定位系统）接收器；常规膝上型和/或掌上型计算机或其他设备，其具有和/或包括射频接收器的常规膝上型和/或掌上型计算机或其他设备。这里所使用的“终端”、“终端设备”可以是便携式、可运输、安装在交通工具（航空、海运和/或陆地）中的，或者适合于和/或配置为在本地运行，和/或以分布形式，运行在地球和/或空间的任何其他位置运行。这里所使用的“终端”、“终端设备”还可以是通信终端、上网终端、音乐/视频播放终端，例如可以是

PDA、MID（Mobile Internet Device，移动互联网设备）和/或具有音乐/视频播放功能的移动电话，也可以是智能电视、机顶盒等设备。

具体请参阅图 1，图 1 为本实施例快速模型的训练方法的基本流程示意图。

如图 1 所示，一种快速模型的训练方法，包括：

S1100、获取预设的训练样本图像；

在对快速模型进行训练之前需要准备训练样本图像，训练样本图像的数量根据对快速模型的具体要求相对应，例如，对于快速模型泛化能力要求较高的应用场景要求的训练样本图像的数量越多，反之，则越少。

对于训练样本图像的获取能够采用公共图像数据库、自有图像数据库或者通过数据爬取的方式从互联网中进行爬取。

快速模型是指与现有的神经网络模型模型相比，需要通过训练确定的自由量（权重）的数量较少的、模型规模较小或者仅仅需要少量的训练样本图像就能够训练至收敛的神经网络模型。

快速模型由于其规模较小，对于计算机处理能力的要求较低，且实现的功能略简单，因此，常使用于（不限于）移动智能终端或者环境因素（需要处理的图像变化较为单一）变化单一的应用场景中。

S1200、将所述训练样本图像输入到预设的辅助训练模型和初始的快速模型中，其中，所述辅助训练模型为预先训练至收敛状态用于提取图像特征向量的神经网络模型，所述快速模型为待训练的神经网络模型；

本实施方式中，为实现快速模型的训练，需要采用辅助训练模型进行训练。且本实施方式中参与训练的辅助训练模型被预先训练至收敛状态，即能够对训练样本图像做出准确分类的神经网络模型。本实施方式中，辅助训练模型能够为已经训练至收敛状态的卷积神经网络模型（CNN），但是辅助训练模型还能够是：深度神经网络模型（DNN）、循环神经网络模型（RNN）或者上述三种网络模型的变形模型。而快速模型也能够为上述三种模型中的任意一种或者变形模型，但是快速模型相比于辅助训练模型规模更小。

需要指出的是快速模型与辅助训练模型的训练方向需一致，或快速模型训练实现的目的是辅助训练模型的功能当中的一种。例如，快速模型训练的目的在于，对黄种人的人脸图像进行识别，辅助训练模型的功能也应该为对人脸图像进行识别，但是相对于快速模型而言，辅助训练模型具有更强大的识别能力，如能够对全人种的人脸图像进行识别，或者不仅仅能够识别人脸图像还能够识别人脸图像表征的用户年龄、性别或者颜值等。

在获取到训练样本图像后，对训练样本图像进行图像处理（将训练样本图像缩放或者剪切成为固定规格的图片）后，将训练样本图像分别输入到辅助训练模型和快速模型中。至此，辅助训练模型和快速模型分别对训练样本图像的特征向量进行提取。

S1300、计算所述辅助训练模型提取的训练样本图像的特征向量与所述快速模型提取的训练样本图像的特征向量之间的特征距离；

分别获取辅助训练模型和快速模型提取的训练样本图像的特征向量。然后计算两个模型特征向量之间的特征距离。其中特征距离包括两个特征向量之间的欧氏距离和/或余弦距离。特征距离的计算是通过损失函数进行的。在一些实施方式中，损失函数既能够计算两个模型特征向量之间的欧氏距离，又能够计算两个模型特征向量之间的余弦距离，并且通过欧氏距离和余弦距离的结合从不同维度，对快速模型的权值进行校正，加速快速模型的训练，提高快速模型的准确度。

由于，辅助训练模型已经预先被训练至收敛状态了，因此，辅助训练模型提取的训练样本图像的特征向量为准确的特征向量，而训练过程中，快速模型输出的特征向量则准确性较低。计算两个模型特征向量之间的距离，其实就是计算快速模型输出的特征向量距标准特征向量之间的距离。

S1400、通过对所述特征距离进行反向传播，以校正所述快速模型中的权重参数。

在计算得到两个模型提取同一训练样本图像的特征向量的特征距离后，将该特征向量通过损失函数进行反向传播，以校正快速模型中的图像过滤器（卷积层）的权重，校正的结果为使快速模型的提取的特征向量向辅助训练模型提取的特征向量靠近。至此，对于快速模型训练的一个环节完成，需要指出的是，步骤 S1100-S1200 仅仅是快速模型训练过程中的一个步骤，由于训练样本图像的数量，步骤 S1100-S1200 的训练是不断循环的，直至快速模型被训练至收敛状态时为止。

上述实施方式通过在对快速模型进行训练时，无需对参与训练的样本图像进行标记，节约了标记所需要的时间和花费的精力，提高了训练的速度。同时，直接计算辅助模型输出的表征样本图像特征的特征向量，与快速模型输出的表征样本图像特征的特征向量之间的距离（欧式距离和/或余弦距离）并进行反向传播，采用这种方法将快速模型的训练转化为一种单纯的回归算法，能够最大限度的缩短训练时间，且能够保证训练完成时快速模型输出的准确率。

在一些实施方式中，快速模型的训练过程并非每次都要进行反向传播的，当两个模型之间的特征距离大于设定的阈值时，才需要进行反向传播。请参阅图 2，图 2 为本实施例通过阈值确定是否进行反向传播的流程示意图。

如图 2 所示，S1400 之前还包括：

S1311、将所述特征距离与预设的第一阈值进行比对；

为了验证两个模型提取的特征向量之间的距离是否满足模型训练的需求，设定一个检验阈值，即第一阈值。第一阈值的设定能够根据对快速模型的准确度要求加以设定。即对于快速模型准确度要求越高，则第一阈值的设定数值越小，反之，对于快速模型的准确度要求越低，则第一阈值的设定数值越大。

S1312、当所述特征距离大于所述第一阈值时，确认对所述特征距离进行反向传播。

通过比较确定特征距离大于第一阈值时，即确定快速模型提取的特征向量与辅助训练模型提取的特征向量之间的距离较大，不符合快速模型训练的需求，需要通过反向传播来调整快速模型的权值，以使再次输入该训练样本图像时，快速模型提取的特征向量与辅助训练模型提取的特征向量之间的距离趋向于变小。当通过比较确定特征距离小于等于第一阈值时，则证明快速模型对于当前输入的训练样本图像的图像理解能力达到了设定要求无需进行反向传播，而是采用下一张训练样本图像对快速模型进行训练。

在本实施方式中，需要指出的是步骤 S1100-S1200 以及步骤 S1311-S1312 是循环进行的。S1400 之后还包括

S1411、反复迭代将所述训练样本图像输入到所述辅助训练模型和所述快速模型中，至所述特征距离小于等于所述第一阈值时，确认所述训练样本图像的训练结束。

当通过比较确定特征距离大于第一阈值时，向同一张训练样本图像再次输入到的辅助训练模型和快速模型中，反复的执行 S1200-S1400 的步骤，直至当通过比较确定特征距离小于等于第一阈值时，对于该训练样本图像的训练结束，还其他训练样本图像继续进行训练。

通过上述方法当快速模型提取的特征向量与辅助训练模型提取的特征向量之间的距离均小于第一阈值时，或者小于第一阈值的比例大于设定值，如 99%或其他比率设定值时，对于快速样本的训练结束。

在一些实施方式中，特征距离包括欧氏距离和/或余弦距离，计算两个模型提取特征向量之间的特征距离，就是计算特征向量之间的欧氏距离和/或余弦距离。请参阅图 3，图 3 为本实施例计算两个模型提取特征向量之间的特征距离的一种流程示意图。

如图 3 所述，S1300 还包括：

S1321、获取所述辅助训练模型提取的所述训练样本图像的第一特征向量和所述快速模型提取的所述训练样本图像的第二特征向量；

分别获取辅助训练模型和快速模型提取的训练样本图像的特征向量。其中，辅助训练模型提取的训练样本图像的特征向量为第一特征向量，快速模型提取的训练样本

图像的特征向量为第二特征向量。

S1322、比对所述第一特征向量和所述第二特征向量的欧氏距离和/或余弦距离。

计算第一特征向量和第二特征向量之间的特征距离。其中特征距离包括两个特征向量之间的欧氏距离和/或余弦距离。特征距离的计算是通过损失函数进行的。在一些实施方式中，损失函数既能够计算两个模型特征向量之间的欧氏距离，又能够计算两个模型特征向量之间的余弦距离，并且通过欧氏距离和余弦距离的结合从不同维度，对快速模型的权值进行校正，加速快速模型的训练，提高快速模型的准确度。

在一些实施方式中，为了使提高快速模型在具体应用环境下的准确性，需要获取具有针对性的训练样本图像对快速模型进行训练。请参阅图 4，图 4 为本实施例根据具体应用场景获取训练样本图像的流程示意图。

如图 4 所示，S1100 之前还包括：

S1011、获取目标用户或者目标场景中需要辨识图像的图像信息，其中，所述图像信息包括各辨识图像的图像类别；

在对快速模型进行训练之前，首先需要收集合适的素材，而为了增强快速模型在实际应用中的准确性，需要对使用快速模型的具体用户或者某个具体场景中的类别图片进行辨识。例如，使用快速模型的用户为资深的宠物爱好者，用户经常上传各种宠物图片进行物种辨识，则收集该用户上传的辨识图像。

其中，每张辨识图像均对应有一个图像信息，图像信息中包括各个辨识图像的图像类别。例如，辨识图像为宠物照片时，图像类型为宠物的物种名称。

S1012、根据所述图像信息中的图像类别对辨识图像进行聚类，以获取所述目标用户或者目标场景偏好辨识的图像类型；

根据获取的各个辨识图像的图像类别，对用户图像进行聚类统计。聚类统计即将辨识图像中图像类型相同的一类聚集为一类。然后根据同一个图像类型中辨识图像的数量确定用户或者应用场景中偏好辨识的图像类型。例如，用户上传各种动物图像上传进行辨识，但是由于用户是狂热的犬科动物爱好者，因此，上传的辨识图像大多数为全科动物的图像，因此，通过统计聚类用户偏好辨识的图像类型为犬科类图像。又或者，用户为旅行家，上传的辨识图像均为旅行途中的动植物图像，通过统计确认用户最近上传的图像均为非洲大陆的动植物图像，则确定用户在一段时间内身处非洲或者对非洲动植物产生较大的兴趣，确定偏好辨识的图像类型为非洲动植物图像。

S1013、根据所述图像类型收集训练快速模型的训练样本图像。

通过图像类型收集训练快速模型的训练样本图像。训练样本图像的获取能够采用公共图像数据库、自有图像数据库或者通过数据爬取的方式从互联网中进行爬取。例

如，通过非洲国家博物馆的数据库，获取非洲动植物样本的图片作为训练样本图像，又或者通过网络爬虫爬取宠物犬的图片作为训练样本图像。

由于，在某种特殊的应用场景中使用时，要求快速模型无需太高的泛化能力，但是对于特定事物图像辨识准确度的要求确很高，通过识别该应用场景中或者用户偏好，来确定收集训练样本图像，对快速模型进行训练，能够使快速模型在针对的领域具有较高的准确性，很好的满足了用户的需求。

在一些实施方式中，训练样本图像均存储在图像数据库中，需要对某个快速模型进行训练时，需要通过检索的方式获取需要的训练样本图像。请参阅图 5，图 5 为本实施例数据库检索获取训练样本图像的一种流程示意图。

如图 5 所示，S1013 之后包括：

S1021、以所述图像类型为限定条件在预设的图像数据库中进行检索；

获取到用户或者应用场景中偏好的图像类型后，以该图像类型为检索条件在预设的图像数据库中进行检索。其中，图像数据库中的所有图像在入库时，均根据图像内容对图像进行标签的设置，例如，是图像为动物时，该图像的标签包括该动物的纲、目、科、属和名称等。

以该图像类型未检索条件进行检索，能够检索到图像标签与该图像类型相同或者相似的图像。

S1022、将检索召回的图像确认为所述训练样本图像。

将通过检索召回的图像确认为对应快速模型的训练样本图像。

在一些实施方式中，为弥补训练样本图像的不足，也为了增强快速模型的稳定性和抗干扰能力，在获取了原始样本图像后，对样本图像进行图像处理，生成派生于原始样本图像的派生样本图像。请参阅图 6，图 6 为本实施例派生样本图像的生成方法流程示意图。

如图 6 所示，S1022 之后还包括：

S1031、对所述原始样本图像进行图像处理，生成派生于所述原始样本图像的派生样本图像；

通过数据库或者通过网络爬虫技术爬取的训练样本图像被定义为原始样本图像。对原始样本图像进行图像处理生成的新的样本图像被定义为派生样本图像。

本实施方式中，对原始样本图像进行图像处理的方法包括（不限于）：对原始样本图像进行图像剪切、图像旋转、或者噪声干扰当中的一种或者多种处理方式进行组合。

由于，原始样本图像与派生样本图像表达的实质内容完全相同，因此，辅助训练模型提取原始样本图像与派生样本图像的特征向量应当完全相同或者相似。因此，需

要进一步的确认图像处理后的派生样本图像,是否形变到模型无法辨识的程度,若是,则该派生样本图像不能作为训练样本。

S1032、提取所述原始样本图像的特征向量和所述派生样本图像的特征向量;

将原始样本图像和派生样本图像先后输入到辅助训练模型中,提取原始样本图像的特征向量和派生样本图像的特征向量。

提取特征向量的模型不局限于辅助训练模型,在一些实施方式中,提取特征向量的模型还能够是已经训练至收敛状态,与快速模型具有相同的训练方向的神经网络模型。

S1033、计算所述原始样本图像的特征向量和所述派生样本图像的特征向量之间的特征差值;

计算得到原始样本图像的特征向量和派生样本图像的特征向量后,计算原始样本图像的特征向量和派生样本图像的特征向量之间的特征差值。其中特征差值计算通过损失函数计算,且特征差值能够是两个特征向量的欧氏距离和/或余弦距离。

S1034、当所述特征差值小于等于预设的第二阈值时,确认所述原始样本图像和派生样本图像为所述训练样本图像。

将计算得到的特征差值与预设的第二阈值进行比对,其中,为了验证两个特征向量之间的距离是否满足模型训练的需求,设定一个检验阈值,即第二阈值。第二阈值的设定能够根据对快速模型的靠干扰要求加以设定。即对于快速模型抗干扰要求越高,则第二阈值的设定数值越小,反之,对于快速模型的抗干扰要求越低,则第二阈值的设定数值越大。

当特征差值小于等于预设的第二阈值时,则表明派生样本图像符合训练的要求,确认原始样本图像和派生样本图像为所述训练样本图像;否则,则确认该派生样本图像不否和训练要求被丢弃。

由于,训练样本图像中的派生样本图像是由原始样本图像进行图像处理得到的,其中包含的图像内容实质相同,且通过特征差值的验证,能够保证通过上述训练样本图像训练至收敛的快速模型,即使在分类的图像存在干扰的情况下也能够准确的提取出正确的特征向量,提高了模型的抗干扰性和稳定性。同时,由于快速模型的训练过程中,训练样本图像无需进行标定,因此,通过计算原始样本图像和派生样本图像之间的特征差值,能够保证快速模型能够准确的提取具有干扰图像的特征向量。

为解决上述技术问题,本申请实施例还提供一种快速模型的训练装置。

具体请参阅图7,图7为本实施例快速模型的训练装置基本结构示意图。

如图7所示,一种快速模型的训练装置,包括:获取模块2100、处理模块2200、

计算模块 2300 和执行模块 2400。其中, 获取模块 2100 用于获取预设的训练样本图像; 处理模块 2200 用于将训练样本图像输入到预设的辅助训练模型和初始的快速模型中, 其中, 辅助训练模型为预先训练至收敛状态用于提取图像特征向量的神经网络模型, 快速模型为待训练的神经网络模型; 计算模块 2300 用于计算辅助训练模型提取的训练样本图像的特征向量与快速模型提取的训练样本图像的特征向量之间的特征距离; 执行模块 2400 用于通过对特征距离进行反向传播, 以校正快速模型中的权重参数。

快速模型的训练装置通过在对快速模型进行训练时, 无需对参与训练的样本图像进行标记, 节约了标记所需要的时间和花费的精力, 提高了训练的速度。同时, 直接计算辅助模型输出的表征样本图像特征的特征向量, 与快速模型输出的表征样本图像特征的特征向量之间的距离 (欧式距离和/或余弦距离) 并进行反向传播, 采用这种方法将快速模型的训练转化为一种单纯的回归算法, 能够最大限度的缩短训练时间, 且能够保证训练完成时快速模型输出的准确率。

在一些实施方式中, 快速模型的训练装置还包括: 第一比对子模块和第一执行子模块。其中, 第一比对子模块用于将特征距离与预设的第一阈值进行比对; 第一执行子模块用于当特征距离大于第一阈值时, 确认对特征距离进行反向传播。

在一些实施方式中, 快速模型的训练装置还包括: 第二执行子模块, 用于反复迭代将训练样本图像输入到辅助训练模型和快速模型中, 至特征距离小于等于第一阈值时, 确认训练样本图像的训练结束。

在一些实施方式中, 特征距离包括欧氏距离和/或余弦距离, 快速模型的训练装置还包括: 第一获取子模块和第二比对子模块。其中, 第一获取子模块用于获取辅助训练模型提取的训练样本图像的第一特征向量和快速模型提取的训练样本图像的第二特征向量; 第二比对子模块用于比对第一特征向量和第二特征向量的欧氏距离和/或余弦距离。

在一些实施方式中, 快速模型的训练装置还包括: 第二获取子模块、第一处理子模块和第三执行子模块。其中, 第二获取子模块用于获取目标用户或者目标场景中需要辨识图像的图像信息, 其中, 图像信息包括各辨识图像的图像类别; 第一处理子模块用于根据图像信息中的图像类别对辨识图像进行聚类, 以获取目标用户或者目标场景偏好辨识的图像类型; 第三执行子模块用于根据图像类型收集训练快速模型的训练样本图像。

在一些实施方式中, 快速模型的训练装置还包括: 第二处理子模块和第一确认子模块。其中, 第二处理子模块用于以图像类型为限定条件在预设的图像数据库中进行检索; 第一确认子模块用于将检索召回的图像确认为训练样本图像。

在一些实施方式中，快速模型的训练装置还包括：第三处理子模块、第一提取子模块、第一计算子模块和第四执行子模块。其中，第三处理子模块用于对原始样本图像进行图像处理，生成派生于原始样本图像的派生样本图像；第一提取子模块用于提取原始样本图像的特征向量和派生样本图像的特征向量；第一计算子模块用于计算原始样本图像的特征向量和派生样本图像的特征向量之间的特征差值；第四执行子模块用于当特征差值小于等于预设的第二阈值时，确认原始样本图像和派生样本图像为训练样本图像。

为解决上述技术问题，本申请实施例还提供计算机设备。具体请参阅图 8，图 8 为本实施例计算机设备基本结构框图。

如图 8 所示，计算机设备的内部结构示意图。该计算机设备包括通过系统总线连接的处理器、非易失性存储介质、存储器和网络接口。其中，该计算机设备的非易失性存储介质存储有操作系统、数据库和计算机可读指令，数据库中可存储有控件信息序列，该计算机可读指令被处理器执行时，可使得处理器实现一种快速模型的训练方法。该计算机设备的处理器用于提供计算和控制能力，支撑整个计算机设备的运行。该计算机设备的存储器中可存储有计算机可读指令，该计算机可读指令被处理器执行时，可使得处理器执行一种快速模型的训练方法。该计算机设备的网络接口用于与终端连接通信。本领域技术人员可以理解，图 8 中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的计算机设备的限定，具体的计算机设备可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

本实施方式中处理器用于执行图 7 中获取模块 2100、处理模块 2200、计算模块 2300 和执行模块 2400 的具体功能，存储器存储有执行上述模块所需的程序代码和各类数据。网络接口用于向用户终端或服务器之间的数据传输。本实施方式中的存储器存储有人脸图像关键点检测装置中执行所有子模块所需的程序代码及数据，服务器能够调用服务器的程序代码及数据执行所有子模块的功能。

计算机设备通过在对快速模型进行训练时，无需对参与训练的样本图像进行标记，节约了标记所需要的时间和花费的精力，提高了训练的速度。同时，直接计算辅助模型输出的表征样本图像特征的特征向量，与快速模型输出的表征样本图像特征的特征向量之间的距离（欧式距离和/或余弦距离）并进行反向传播，采用这种方法将快速模型的训练转化为一种单纯的回归算法，能够最大限度的缩短训练时间，且能够保证训练完成时快速模型输出的准确率。

本申请还提供一种存储有计算机可读指令的存储介质，计算机可读指令被一个或

多个处理器执行时，使得一个或多个处理器执行上述任一实施例快速模型的训练方法的步骤。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，该计算机程序可存储于一计算机可读取存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，前述的存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory, ROM）等非易失性存储介质，或随机存储记忆体（Random Access Memory, RAM）等。

权 利 要 求 书

1.一种快速模型的训练方法，包括：

获取预设的训练样本图像；

将所述训练样本图像输入到预设的辅助训练模型和初始的快速模型中，其中，所述辅助训练模型为预先训练至收敛状态用于提取图像特征向量的神经网络模型，所述快速模型为待训练的神经网络模型；

计算所述辅助训练模型提取的训练样本图像的特征向量与所述快速模型提取的训练样本图像的特征向量之间的特征距离；

通过对所述特征距离进行反向传播，以校正所述快速模型中的权重参数。

2.根据权利要求1所述的快速模型的训练方法，所述通过对所述特征距离进行反向传播，以校正所述快速模型中的权重参数之前还包括：

将所述特征距离与预设的第一阈值进行比对；

当所述特征距离大于所述第一阈值时，确认对所述特征距离进行反向传播。

3.根据权利要求2所述的快速模型的训练方法，所述通过对所述特征距离进行反向传播，以校正所述快速模型中的权重参数之后还包括：

反复迭代将所述训练样本图像输入到所述辅助训练模型和所述快速模型中，至所述特征距离小于等于所述第一阈值时，确认所述训练样本图像的训练结束。

4.根据权利要求1所述的快速模型的训练方法，所述特征距离包括欧氏距离和/或余弦距离，所述计算所述辅助训练模型提取的训练样本图像的特征向量与所述快速模型提取的训练样本图像的特征向量之间的特征距离包括：

获取所述辅助训练模型提取的所述训练样本图像的第一特征向量和所述快速模型提取的所述训练样本图像的第二特征向量；

比对所述第一特征向量和所述第二特征向量的欧氏距离和/或余弦距离。

5.根据权利要求1所述的快速模型的训练方法，所述获取预设的训练样本图像之前还包括：

获取目标用户或者目标场景中需要辨识图像的图像信息，其中，所述图像信息包括各辨识图像的图像类别；

根据所述图像信息中的图像类别对辨识图像进行聚类，以获取所述目标用户或者目标场景偏好辨识的图像类型；

根据所述图像类型收集训练快速模型的训练样本图像。

6.根据权利要求5所述的快速模型的训练方法,所述根据所述图像类型收集训练快速模型的训练样本图像包括:

以所述图像类型为限定条件在预设的图像数据库中进行检索;

将检索召回的图像确认为所述训练样本图像。

7.根据权利要求6所述的快速模型的训练方法,所述训练样本图像包括原始样本图像和派生样本图像,所述将检索召回的图像确认为所述训练样本图像之后还包括:

对所述原始样本图像进行图像处理,生成派生于所述原始样本图像的派生样本图像;

提取所述原始样本图像的特征向量和所述派生样本图像的特征向量;

计算所述原始样本图像的特征向量和所述派生样本图像的特征向量之间的特征差值;

当所述特征差值小于等于预设的第二阈值时,确认所述原始样本图像和派生样本图像为所述训练样本图像。

8.一种快速模型的训练装置,包括:

获取模块,用于获取预设的训练样本图像;

处理模块,用于将所述训练样本图像输入到预设的辅助训练模型和初始的快速模型中,其中,所述辅助训练模型为预先训练至收敛状态用于提取图像特征向量的神经网络模型,所述快速模型为待训练的神经网络模型;

计算模块,用于计算所述辅助训练模型提取的训练样本图像的特征向量与所述快速模型提取的训练样本图像的特征向量之间的特征距离;

执行模块,用于通过对所述特征距离进行反向传播,以校正所述快速模型中的权重参数。

9.一种计算机设备,包括存储器和处理器,所述存储器中存储有计算机可读指令,所述计算机可读指令被所述处理器执行时,使得所述处理器执行一种快速模型的训练方法,所述方法包括如下步骤:

获取预设的训练样本图像;

将所述训练样本图像输入到预设的辅助训练模型和初始的快速模型中,其中,所述辅助训练模型为预先训练至收敛状态用于提取图像特征向量的神经网络模型,所述快速模型为待训练的神经网络模型;

计算所述辅助训练模型提取的训练样本图像的特征向量与所述快速模型提取的训练样本图像的特征向量之间的特征距离;

通过对所述特征距离进行反向传播,以校正所述快速模型中的权重参数。

10. 根据权利要求 9 所述的计算机设备, 所述通过对所述特征距离进行反向传播, 以校正所述快速模型中的权重参数之前还包括:

将所述特征距离与预设的第一阈值进行比对;

当所述特征距离大于所述第一阈值时, 确认对所述特征距离进行反向传播。

11. 根据权利要求 10 所述的计算机设备, 所述通过对所述特征距离进行反向传播, 以校正所述快速模型中的权重参数之后还包括:

反复迭代将所述训练样本图像输入到所述辅助训练模型和所述快速模型中, 至所述特征距离小于等于所述第一阈值时, 确认所述训练样本图像的训练结束。

12. 根据权利要求 9 所述的计算机设备, 所述特征距离包括欧氏距离和/或余弦距离, 所述计算所述辅助训练模型提取的训练样本图像的特征向量与所述快速模型提取的训练样本图像的特征向量之间的特征距离包括:

获取所述辅助训练模型提取的所述训练样本图像的第一特征向量和所述快速模型提取的所述训练样本图像的第二特征向量;

比对所述第一特征向量和所述第二特征向量的欧氏距离和/或余弦距离。

13. 根据权利要求 9 所述的计算机设备, 所述获取预设的训练样本图像之前还包括:

获取目标用户或者目标场景中需要辨识图像的图像信息, 其中, 所述图像信息包括各辨识图像的图像类别;

根据所述图像信息中的图像类别对辨识图像进行聚类, 以获取所述目标用户或者目标场景偏好辨识的图像类型;

根据所述图像类型收集训练快速模型的训练样本图像。

14. 根据权利要求 13 所述的计算机设备, 所述根据所述图像类型收集训练快速模型的训练样本图像包括:

以所述图像类型为限定条件在预设的图像数据库中进行检索;

将检索召回的图像确认为所述训练样本图像。

15. 根据权利要求 14 所述的计算机设备, 所述训练样本图像包括原始样本图像和派生样本图像, 所述将检索召回的图像确认为所述训练样本图像之后还包括:

对所述原始样本图像进行图像处理, 生成派生于所述原始样本图像的派生样本图像;

提取所述原始样本图像的特征向量和所述派生样本图像的特征向量;

计算所述原始样本图像的特征向量和所述派生样本图像的特征向量之间的特征差值;

当所述特征差值小于等于预设的第二阈值时, 确认所述原始样本图像和派生样本

图像为所述训练样本图像。

16. 一种存储有计算机可读指令的非易失性存储介质, 所述计算机可读指令被一个或多个处理器执行时, 使得一个或多个处理器执行一种快速模型的训练方法, 所述方法包括以下步骤:

获取预设的训练样本图像;

将所述训练样本图像输入到预设的辅助训练模型和初始的快速模型中, 其中, 所述辅助训练模型为预先训练至收敛状态用于提取图像特征向量的神经网络模型, 所述快速模型为待训练的神经网络模型;

计算所述辅助训练模型提取的训练样本图像的特征向量与所述快速模型提取的训练样本图像的特征向量之间的特征距离;

通过对所述特征距离进行反向传播, 以校正所述快速模型中的权重参数。

17. 根据权利要求 16 所述的非易失性存储介质, 所述通过对所述特征距离进行反向传播, 以校正所述快速模型中的权重参数之前还包括:

将所述特征距离与预设的第一阈值进行比对;

当所述特征距离大于所述第一阈值时, 确认对所述特征距离进行反向传播。

18. 根据权利要求 17 所述的非易失性存储介质, 所述通过对所述特征距离进行反向传播, 以校正所述快速模型中的权重参数之后还包括:

反复迭代将所述训练样本图像输入到所述辅助训练模型和所述快速模型中, 至所述特征距离小于等于所述第一阈值时, 确认所述训练样本图像的训练结束。

19. 根据权利要求 16 所述的非易失性存储介质, 所述特征距离包括欧氏距离和/或余弦距离, 所述计算所述辅助训练模型提取的训练样本图像的特征向量与所述快速模型提取的训练样本图像的特征向量之间的特征距离包括:

获取所述辅助训练模型提取的所述训练样本图像的第一特征向量和所述快速模型提取的所述训练样本图像的第二特征向量;

比对所述第一特征向量和所述第二特征向量的欧氏距离和/或余弦距离。

20. 根据权利要求 16 所述的非易失性存储介质, 所述获取预设的训练样本图像之前还包括:

获取目标用户或者目标场景中需要辨识图像的图像信息, 其中, 所述图像信息包括各辨识图像的图像类别;

根据所述图像信息中的图像类别对辨识图像进行聚类, 以获取所述目标用户或者目标场景偏好辨识的图像类型;

根据所述图像类型收集训练快速模型的训练样本图像。

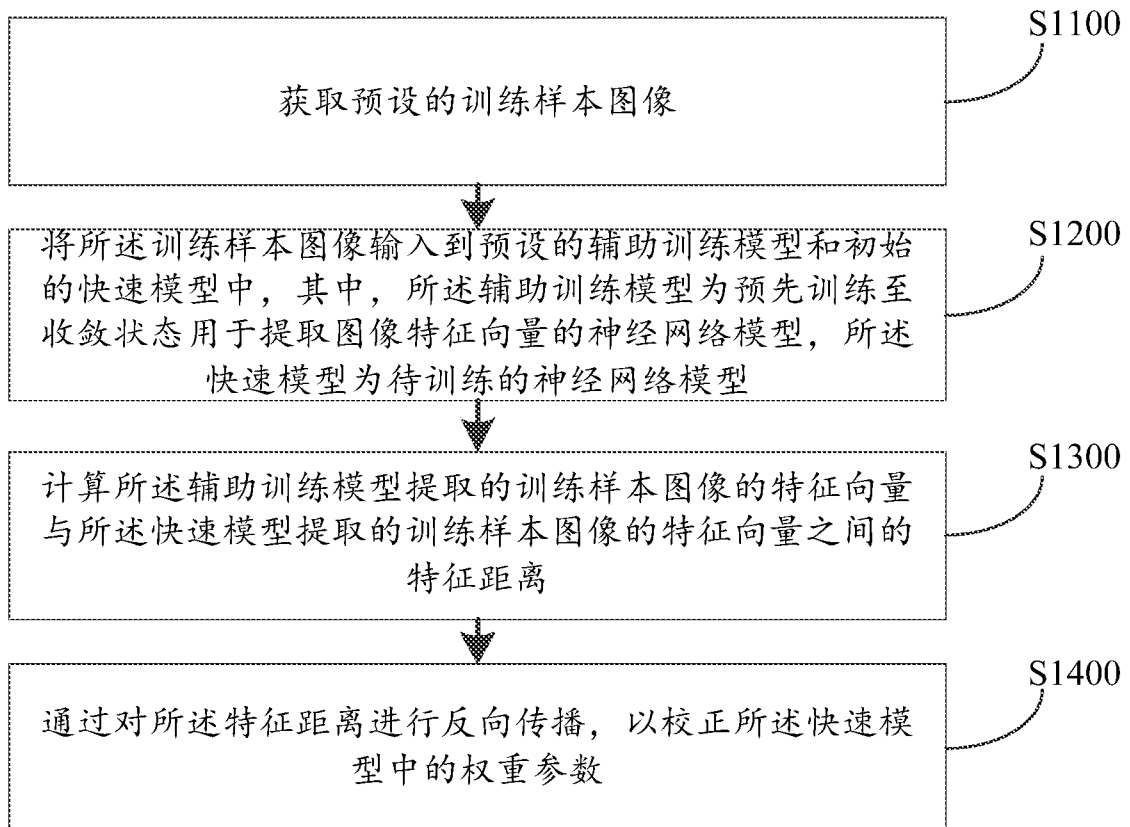


图 1

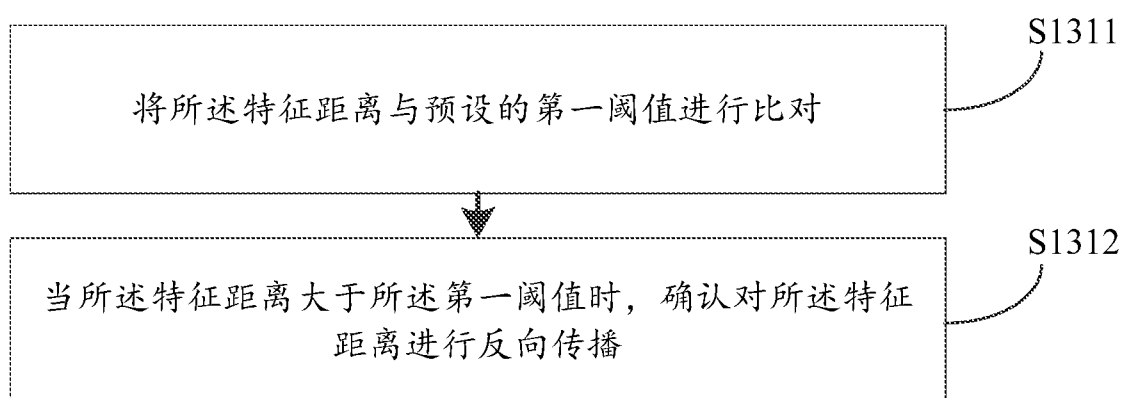


图 2

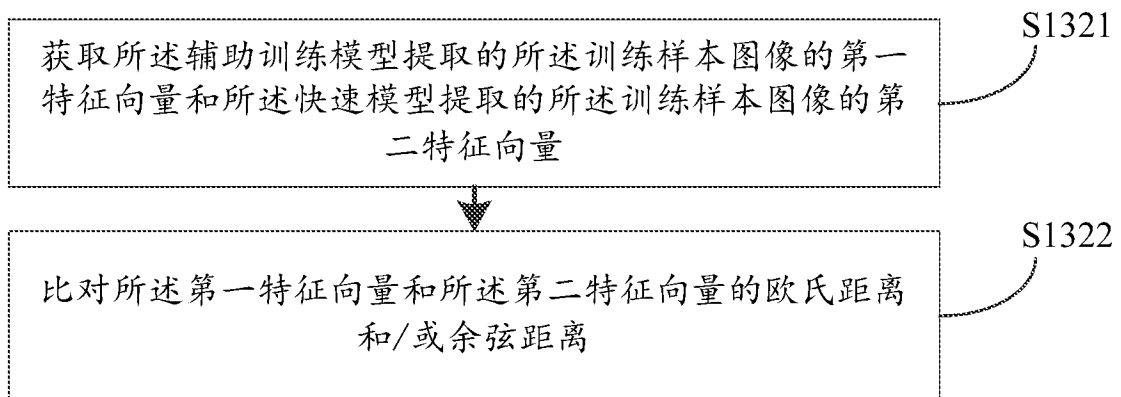


图 3

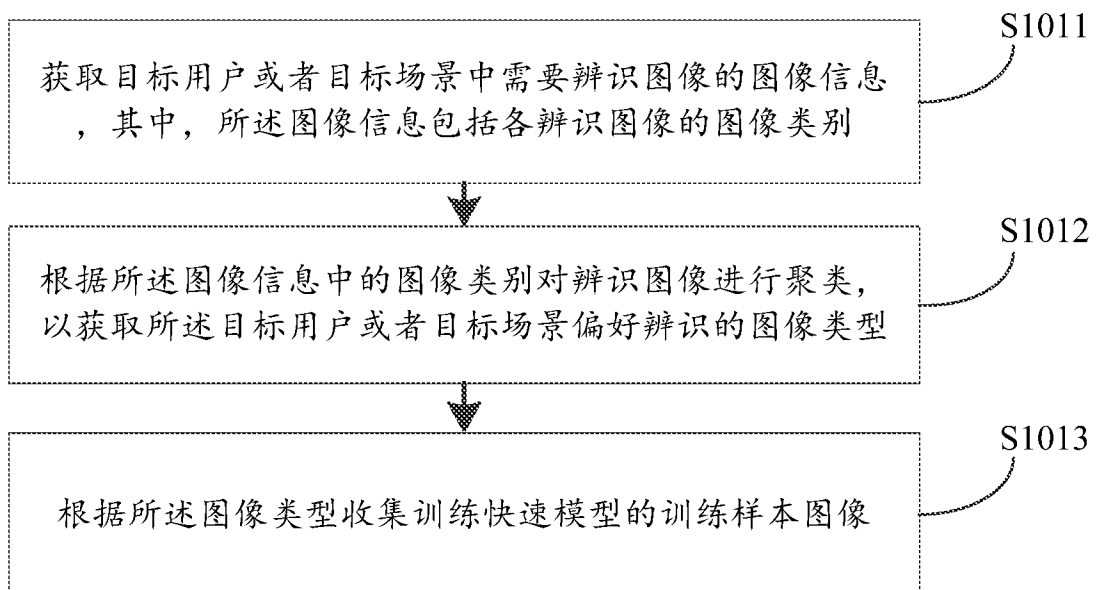


图 4

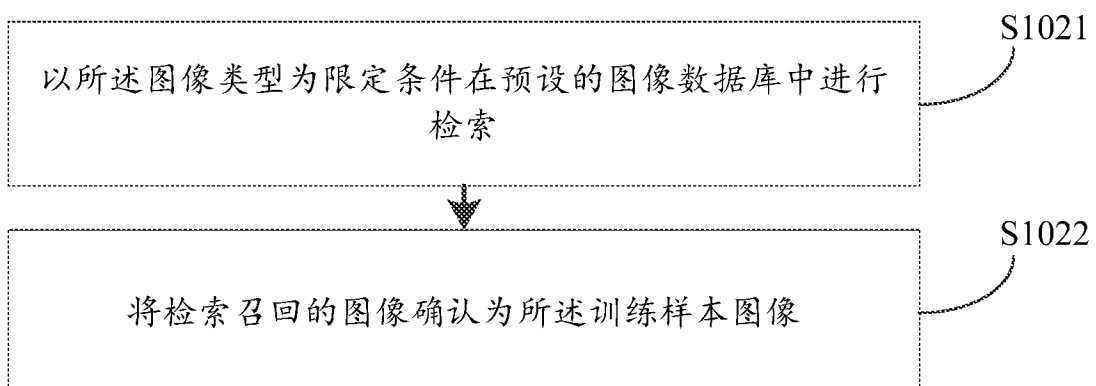


图 5

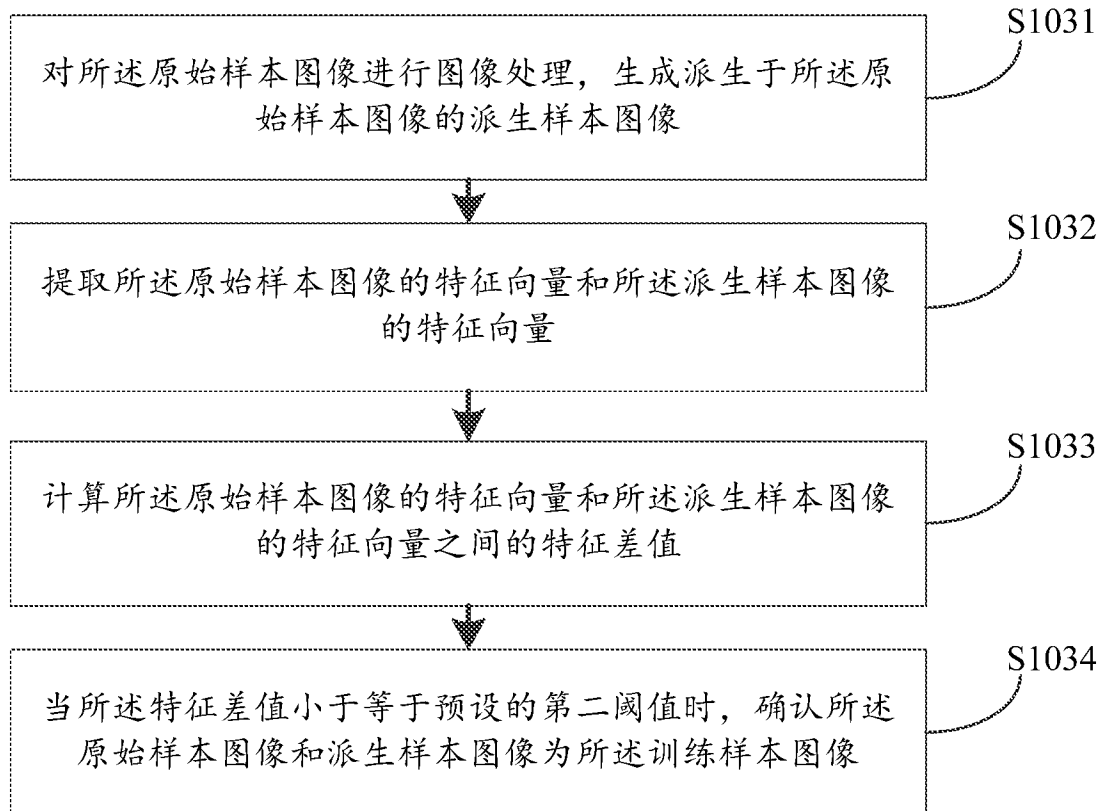


图 6

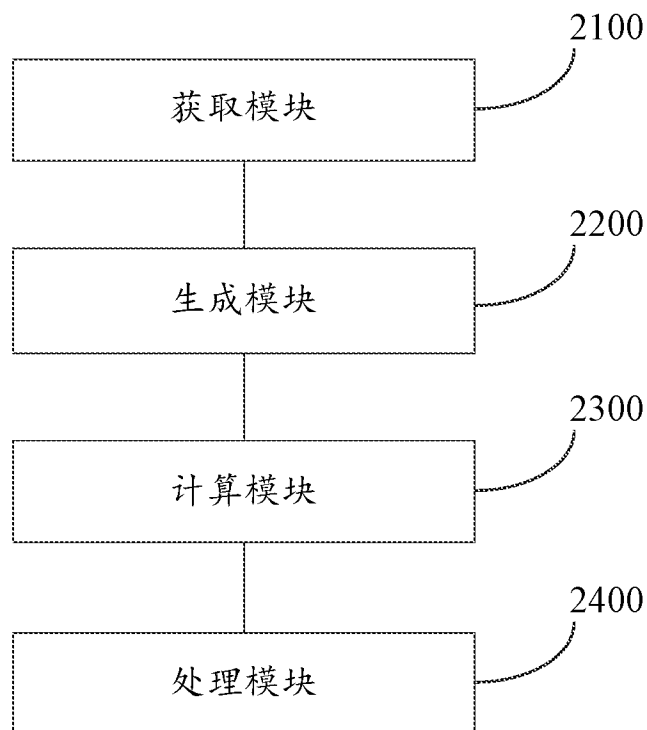


图 7

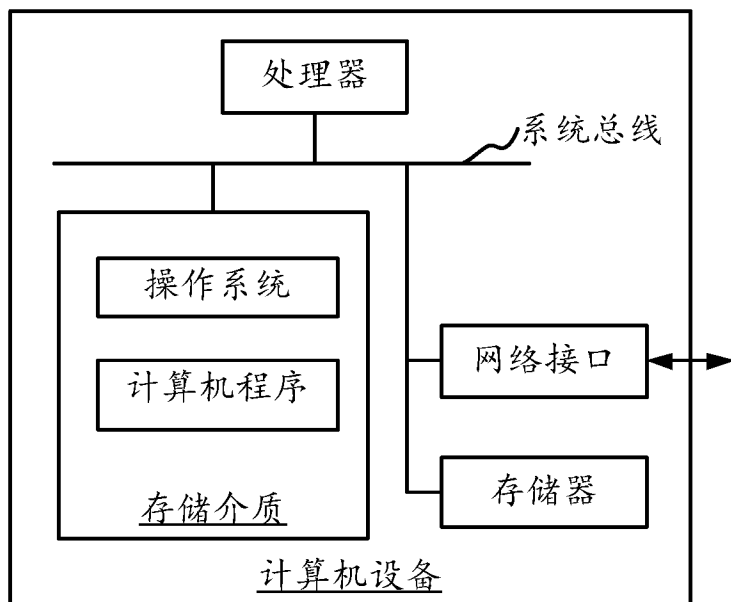


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/125592

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/62(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; TWABS; TWTXT; VEN; USTXT; EPTXT; WOTXT; CNKI: 模型, 神经网络, 教师网络, 学生网络, 训练, 距离, 相似, 反向传播, 后向传播, 权重, 权值, model, neural network, teacher network, student network, train, distance, similarity, back propagation, BP, weight

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 107247989 A (BEIJING TUSIMPLE FUTURE TECHNOLOGY CO., LTD.) 13 October 2017 (2017-10-13) description, paragraphs 37-69	1-20
X	CN 108133238 A (SIC YOUE DATA CO., LTD.) 08 June 2018 (2018-06-08) description, paragraphs 40-66	1-20
A	CN 108182394 A (ZHEJIANG DAHUA TECHNOLOGY CO., LTD.) 19 June 2018 (2018-06-19) entire document	1-20
A	US 2015170020 A1 (AMAZON TECH INC.) 18 June 2015 (2015-06-18) entire document	1-20



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

02 August 2019

Date of mailing of the international search report

22 August 2019

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC (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/CN2018/125592

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	107247989	A	13 October 2017	US	2018365564	A1	20 December 2018
CN	108133238	A	08 June 2018	None			
CN	108182394	A	19 June 2018	None			
US	2015170020	A1	18 June 2015	WO	2015089148	A3	24 September 2015
				US	9400955	B2	26 July 2016
				WO	2015089148	A2	18 June 2015

国际检索报告

国际申请号

PCT/CN2018/125592

A. 主题的分类

G06K 9/62 (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; TWABS; TWTXT; VEN; USTXT; EPTXT; WOTXT; CNKI: 模型, 神经网络, 教师网络, 学生网络, 训练, 距离, 相似, 反向传播, 后向传播, 权重, 权值, model, neural network, teacher network, student network, train, distance, similarity, back propagation, BP, weight

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 107247989 A (北京图森未来科技有限公司) 2017年 10月 13日 (2017 - 10 - 13) 说明书第37-69段	1-20
X	CN 108133238 A (国信优易数据有限公司) 2018年 6月 8日 (2018 - 06 - 08) 说明书第40-66段	1-20
A	CN 108182394 A (浙江大华技术股份有限公司) 2018年 6月 19日 (2018 - 06 - 19) 全文	1-20
A	US 2015170020 A1 (AMAZON TECH INC) 2015年 6月 18日 (2015 - 06 - 18) 全文	1-20

☐ 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

2019年 8月 2日

国际检索报告邮寄日期

2019年 8月 22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

受权官员

彭玉静

电话号码 86-(20)-28958030

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/125592

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	107247989	A	2017年 10月 13日	US	2018365564	A1	2018年 12月 20日
CN	108133238	A	2018年 6月 8日	无			
CN	108182394	A	2018年 6月 19日	无			
US	2015170020	A1	2015年 6月 18日	WO	2015089148	A3	2015年 9月 24日
				US	9400955	B2	2016年 7月 26日
				WO	2015089148	A2	2015年 6月 18日