

## 图像识别模型生成方法、装置、计算机设备和存储介质

### 相关申请交叉引用

本申请要求 2020 年 08 月 25 日递交的、标题为“图像识别模型生成方法、装置、计算机设备和存储介质”、申请号为 2020108629110 的中国申请，其公开内容通过引用全部结合在本申请中。

### 技术领域

本申请涉及一种图像识别模型生成方法、装置、计算机设备和存储介质。

### 背景技术

深度学习中，图像识别技术已经取得了巨大的进展。但是这些进展的取得离不开大规模的数据集，例如 ImageNet、COCO 等。通常情况下，这些大规模的数据集是类别均衡的；但在现实场景下，我们所能获得的数据通常是遵循长尾分布，即小部分类别包含的图像数据很多，而大部分类别包含的图像数据较少。

### 发明内容

根据多个实施例，本申请第一方面提供一种图像识别模型生成方法，所述方法包括：

获取样本图像集；所述样本图像集包括图像数量依次递减的多个样本图像子集，所述多个样本图像子集均包含相同数量的图像类别；

根据所述样本图像集，对待训练的图像识别模型进行训练，得到所述待训练的图像识别模型的损失值；所述待训练的图像识别模型包括多个分支神经网络，每个分支神经网络用于对相应的图像进行识别；所述损失值包括目标分类损失值以及与各个所述分支神经网络对应的分类损失值，所述目标分类损失值为所述待训练的图像识别模型针对所述样本图像集的损失值，所述分类损失值为相应的分支神经网络针对与所述分支神经网络对应的样本图像子集的损失值；以及

根据所述损失值调整所述待训练的图像识别模型的模型参数，直至所述损失值低于预设阈值时，将所述待训练的图像识别模型作为训练完成的图像识别模型。

根据多个实施例，本申请第二方面提供一种图像识别模型生成装置，所述装置包括：

获取模块，用于获取样本图像集；所述样本图像集包括图像数量依次递减的多个样本图像子集，所述多个样本图像子集均包含相同数量的图像类别；

训练模块，用于根据所述样本图像集，对待训练的图像识别模型进行训练，得到所述待训练的图像识别模型的损失值；所述待训练的图像识别模型包括多个分支神经网络，每个分支神经网络用于对相应的图像进行识别；所述损失值包括目标分类损失值以及与各所述分支神经网络对应的分类损失值，所述目标分类损失值为所述待训练的图像识别模型针对所述样本图像集的损失值，所述分类损失值为相应的分支神经网络针对与所述分支神经网络对应的样本图像子集的损失值；以及

调整模块，用于根据所述损失值调整所述待训练的图像识别模型的模型参数，直至所述损失值低于预设阈值时，将所述待训练的图像识别模型作为训练完成的图像识别模型。

根据多个实施例，本申请第三方面提供一种计算机设备，包括存储器和处理器，所述存储器存储有计算机程序，所述处理器执行所述计算机程序时实现以下步骤：

获取样本图像集；所述样本图像集包括图像数量依次递减的多个样本图像子集，所述多个样本图像子集均包含相同数量的图像类别；

根据所述样本图像集，对待训练的图像识别模型进行训练，得到所述待训练的图像识别模型的损失值；所述待训练的图像识别模型包括多个分支神经网络，每个分支神经网络用于对相应的图像进行识别；所述损失值包括目标分类损失值以及与各所述分支神经网络对应的分类损失值，所述目标分类损失值为所述待训练的图像识别模型针对所述样本图像集的损失值，所述分类损失值为相应的分支神经网络针对与所述分支神经网络对应的样本图像子集的损失值；以及

根据所述损失值调整所述待训练的图像识别模型的模型参数，直至所述损失值低于预设阈值时，将所述待训练的图像识别模型作为训练完成的图像识别模型。

根据多个实施例，本申请第四方面提供一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现以下步骤：

获取样本图像集；所述样本图像集包括图像数量依次递减的多个样本图像子集，所述多个样本图像子集均包含相同数量的图像类别；

根据所述样本图像集，对待训练的图像识别模型进行训练，得到所述待训练的图像识别模型的损失值；所述待训练的图像识别模型包括多个分支神经网络，每个分支神经网络用于对相应的图像进行识别；所述损失值包括目标分类损失值以及与各所述分支神经网络

络对应的分类损失值，所述目标分类损失值为所述待训练的图像识别模型针对所述样本图像集的损失值，所述分类损失值为相应的分支神经网络针对与所述分支神经网络对应的样本图像子集的损失值；以及

根据所述损失值调整所述待训练的图像识别模型的模型参数，直至所述损失值低于预设阈值时，将所述待训练的图像识别模型作为训练完成的图像识别模型。

本申请的一个或多个实施例的细节在下面的附图和描述中提出。本申请的其它特征和优点将从说明书、附图以及权利要求书变得明显。

## 附图说明

为了更清楚地说明本申请实施例或传统技术中的技术方案，下面将对实施例或传统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他的附图。

图 1 为一个实施例中图像识别模型生成方法的应用环境图。

图 2 为一个实施例中图像识别模型生成方法的流程示意图。

图 3 为一个实施例中分支神经网络的结构示意图。

图 4 为一个实施例中对待训练的图像识别模型进行训练得到损失值步骤的流程示意图。

图 5 为一个实施例中确定待训练的图像识别模型的损失值步骤的流程示意图。

图 6 为一个实施例中得到样本图像子集以及样本图像集方法的流程示意图。

图 7 为一个实施例中图像识别模型生成装置的结构框图。

图 8 为一个实施例中计算机设备的内部结构图。

## 具体实施方式

利用这种符合长尾分布的数据对神经网络进行训练，得到的结果通常是神经网络能够很好对小部分包含图像数据较多的类别进行识别，而识别包含的图像数据较少的大部分类别的准确性较差；可见，在图像识别模型生成中若忽略这种长尾分布特性会使得图像识别模型的性能在实际使用中大打折扣。因此，通过现有的图像识别模型生成方法，得到的图像识别模型的识别效果还较差。

为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本

申请进行进一步详细说明。应当理解，此处描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

本申请提供的图像识别模型生成方法，可以应用于如图 1 所示的应用环境中。其中，终端 11 通过网络与服务器 12 进行通信。服务器 12 通过网络获取终端 11 发送的样本图像集；样本图像集包括图像数量依次递减的多个样本图像子集，多个样本图像子集均包含相同数量的图像类别；服务器 12 根据样本图像集，对待训练的图像识别模型进行训练，得到待训练的图像识别模型的损失值；待训练的图像识别模型包括多个分支神经网络，每个分支神经网络用于对相应的图像进行识别；损失值包括目标分类损失值以及与各个分支神经网络对应的分类损失值，目标分类损失值为待训练的图像识别模型针对样本图像集的损失值，分类损失值为相应的分支神经网络针对与分支神经网络对应的样本图像子集的损失值；服务器 12 根据损失值调整待训练的图像识别模型的模型参数，直至损失值低于预设阈值时，将待训练的图像识别模型作为训练完成的图像识别模型。终端 11 能够将待识别的图像发送至服务器 12，并得到服务器 12 返回的识别结果。

其中，终端 11 可以但不限于各种个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备，服务器 12 可以用独立的服务器或者是多个服务器组成的服务器集群来实现。

在一个实施例中，如图 2 所示，提供了一种图像识别模型生成方法，以该方法应用于图 1 中的服务器 12 为例进行说明，包括以下步骤：

步骤 21，获取样本图像集；样本图像集包括图像数量依次递减的多个样本图像子集，多个样本图像子集均包含相同数量的图像类别。

其中，样本图像集为包含所有样本图像的数据集，由多个样本图像子集组成，每个样本图像子集都包含一个或多个图像类别的样本图像，且各个样本图像子集包含的图像类别不同；另外，样本图像子集包含的图像总数不相同，且成依次递减的数量趋势。

例如，样本图像中图像类别 A 有 100 张，图像类别 B 有 80 张，图像类别 C 有 60 张，图像类别 D 有 40 张，图像类别 E 有 20 张，图像类别 F 有 10 张；则图像类别 A、B 可以构成一个包含有 180 张样本图像的样本图像子集，图像类别 C、D 可以构成一个包含有 100 张样本图像的样本图像子集，图像类别 E、F 可以构成一个包含有 30 张样本图像的样本图像子集。可见，三个样本图像子集图像数量依次递减，且包含相同数量的图像类别。

具体地，服务器可以直接从终端处获取包括图像数量依次递减的多个样本图像子集的样本图像集；也可以从终端处获取大量样本图像，根据样本图像对应的图像类型将样本图

像进行分类处理,得到包括图像数量依次递减的多个样本图像子集的样本图像集。样本图像集可以由符合长尾分布特性的样本图像(即小部分图像类别的图像数量较多,而大部分图像类别的图像数量较少)构成,也可以由符合正态分布特性的样本图像构成,此处对样本图像集中样本图像的类别分布特性不做限定。

本步骤通过获取包括图像数量依次递减的多个样本图像子集的样本图像集,实现了对样本图像的预处理,使得样本图像按照图像类别进行排序并处于不同的样本图像子集,便于后续分支神经网络进行特征学习,使得训练过程中图像数量较少的图像类别也能得到充分训练,避免了传统神经网络训练过程中对于长尾数据的忽视,提升了图像识别模型生成的效果。

步骤 22,根据样本图像集,对待训练的图像识别模型进行训练,得到待训练的图像识别模型的损失值;待训练的图像识别模型包括多个分支神经网络,每个分支神经网络用于对相应的图像进行识别;损失值包括目标分类损失值以及与各个分支神经网络对应的分类损失值,目标分类损失值为待训练的图像识别模型针对样本图像集的损失值,分类损失值为相应的分支神经网络针对与分支神经网络对应的样本图像子集的损失值。

具体地,通过  $1 \times 1$  的卷积可以实现分支神经网络的构建,因此分支神经网络构建的过程仅需很少的额外参数即可完成。由于在待训练的图像识别模型中构建了多个分支神经网络,分支神经网络能够将待训练的图像识别模型的参数分为两部分,一部分是用于提取样本图像公共特征的共享参数,另一部分是在共享参数基础上,进一步用于提取与分支神经网络对应的样本图像子集的样本图像的独享参数;其中,独享参数即是分支神经网络中的相应参数。

分支神经网络的构建好后,确定分支神经网络与样本图像子集的对应关系,对应关系可以根据分支神经网络的个数以及样本图像子集进行确定。较为常用的是三个分支神经网络与三个样本图像子集,且规定其中一个分支神经网络与三个样本图像子集对应,第二个分支神经网络与三个样本图像子集中的后两个样本图像子集对应,第三个分支神经网络与三个样本图像子集中最后一个样本图像子集(图像数量最少的样本图像子集)对应。

例如,一样本图像集中包括三个样本图像子集,分别为 head classes(头部数据,简称 h),medium classes(中部数据,简称 m)以及 tail classes(尾部数据,简称 t);head classes 中包括图像数量最多的前  $1/3$  的图像类别,medium classes 中包括图像数量处于中间  $1/3$  的图像类别,tail classes 中包括图像数量最少的后  $1/3$  的图像类别。通过  $1 \times 1$  卷积构建如图 3 所示的三个分支神经网络  $N_{h+m+t}$ 、 $N_{m+t}$  以及  $N_t$ ;其中, $N_{h+m+t}$  与全部样本图像

子集对应，用于对全部样本图像子集中的图像类别进行分类； $N_{m+t}$ 与二个样本图像子集对应，用于对图像数量相对较少 **medium classes** 以及 **tail classes** 样本图像子集中的图像类别进行分类； $N_t$ 与一个样本图像子集对应，用于对图像数量最少的 **tail classes** 样本图像子集中的图像类别进行分类；可见三个分支神经网络  $N_{h+m+t}$ 、 $N_{m+t}$  以及  $N_t$  都能够通过自身的共享参数主导对应的样本图像子集中图像类别的学习；且图像数量较少的 **tail classes** 与三个分支神经网络均存在对应关系，而数量较多的 **head classes** 只与一个分支神经网络均存在对应关系，这在一定程度上实现了对长尾数据的利用程度，使得不同图像数量的图像类别在训练时趋于平衡。

待训练的图像识别模型的损失值包括分类损失值以及目标分类损失值，其中分类损失值是分支神经网络针对与分支神经网络对应的样本图像子集的损失值；目标分类损失值是待训练的图像识别模型针对样本图像集的损失值，可以调整。根据多个分类损失值以及目标分类损失值，可以得到对待训练的图像识别模型进行训练的损失值，判断整个图像识别模型的训练程度。

分类损失值即是分别与分支神经网络对应的样本图像子集的损失值，即分支神经网络  $N_{h+m+t}$  对应于 **head classes**、**medium classes** 以及 **tail classes** 样本图像子集的损失值，或者  $N_t$  对应于 **tail classes** 样本图像子集的损失值。目标分类损失值是待训练的图像识别模型整体输出的图像类别对应于整个样本图像集得到的损失值，即多个分支神经网络识别样本图像集后输出的图像类别进行融合后，得到的图像类别对应于样本图像集的损失值。分类损失值和目标分类损失值的区别即是在计算损失值时考量的对象不同，分类损失值是将各个分支神经网络输出的图像类别与相对应的样本图像子集的样本图像的实际图像类别进行比较后得到的损失值；而目标分类损失值是将整个待训练的图像识别模型输出的图像类别（即多个分支神经网络输出的图像类别的融合结果）与样本图像集中样本图像的实际类别进行比较后得到的损失值。

本步骤通过分支神经网络对相应的图像进行识别，得到目标分类损失值以及与各个分支神经网络对应的分类损失值，对图像识别模型进行训练，得到待训练的图像识别模型的损失值，使得训练过程中图像数量较少的图像类别也能得到充分训练，避免了传统神经网络训练过程中对于长尾数据的忽视，提升了图像识别模型生成的效果。

步骤 23，根据损失值调整待训练的图像识别模型的模型参数，直至损失值低于预设阈值时，将待训练的图像识别模型作为训练完成的图像识别模型。

具体地，服务器根据计算得到的损失值反向调整待训练的图像识别模型中包括但不

限于卷积层、池化层、归一化层等的各项参数，例如权重和偏置；在正常情况下，经过多次训练迭代后各项损失值会逐渐减小，且趋近于一个定值。预设阈值可以设置在该定值附近，当损失值低于预设阈值时，即可判断图像识别模型训练完成。

本步骤通过损失值不断对图像识别模型中的参数做出调整，根据损失值与预设阈值的差距判断出图像识别模型的训练程度，直到图像识别模型计算出的损失值低于预设阈值时，可以判断图像识别模型训练完成；提升了图像识别模型生成的效果。

上述图像识别模型生成方法，包括：获取样本图像集；样本图像集包括图像数量依次递减的多个样本图像子集，多个样本图像子集均包含相同数量的图像类别；根据样本图像集，对待训练的图像识别模型进行训练，得到待训练的图像识别模型的损失值；待训练的图像识别模型包括多个分支神经网络，每个分支神经网络用于对相应的图像进行识别；损失值包括目标分类损失值以及与各个分支神经网络对应的分类损失值，目标分类损失值为待训练的图像识别模型针对样本图像集的损失值，分类损失值为相应的分支神经网络针对与分支神经网络对应的样本图像子集的损失值；以及根据损失值调整待训练的图像识别模型的模型参数，直至损失值低于预设阈值时，将待训练的图像识别模型作为训练完成的图像识别模型。本申请通过设置图像数量依次递减的多个样本图像子集，以及对相应样本图像子集的图像进行识别的分支神经网络，使得训练过程中图像数量较少的图像类别也能得到充分训练，避免了传统神经网络训练过程中对于长尾数据的忽视，提升了图像识别模型生成的效果。

在一个实施例中，如图 4 所示，上述步骤 22，根据样本图像集，对待训练的图像识别模型进行训练，得到待训练的图像识别模型的损失值，包括：

步骤 41，对样本图像集中的多个样本图像子集进行均匀采样，得到样本图像输入序列；

步骤 42，根据样本图像输入序列，将样本图像输入待训练的图像识别模型，得到样本图像的图像类别；

步骤 43，根据样本图像的图像类别以及对应的实际图像类别，确定待训练的图像识别模型的损失值。

具体地，服务器可以对样本图像集中的多个样本图像子集进行均匀采样，得到 mini-batch 数据；将 mini-batch 数据作为样本图像输入序列输入待训练的图像识别模型进行训练，得到图像识别模型输出的样本图像的图像类别；获取样本图像的实际图像类别，将样本图像的图像类别以及实际图像类别输入预设的损失函数，计算得到图像识别模型的

损失值。

本实施例通过均匀采样，使得样本图像输入序列中各个图像类别的样本图像较为均衡；进一步使得确定出的待训练的图像识别模型的损失值更加准确，提升了图像识别模型生成的效果。

在一个实施例中，待训练的图像识别模型中还包括基础神经网络，基础神经网络与分支神经网络相连接；

上述步骤 42，根据样本图像输入序列，将样本图像输入待训练的图像识别模型，得到样本图像的图像类别，包括：将样本图像输入待训练的图像识别模型，使得基础神经网络获取样本图像的第一图像特征，以及使得分支神经网络根据第一图像特征得到样本图像的第二图像特征，并根据第二图像特征确定样本图像集中样本图像的图像类别。

具体地，基础神经网络用于提取样本图像集中样本图像的特征信息，即提取样本图像集中所有图像类别的公共特征作为第一图像特征；分支神经网络获取基础神经网络提取出的第一图像特征，并进行再次提取，得到并输出第二图像特征。分支神经网络输出的第二图像特征通过分类器并进行融合后得到样本图像的图像类别。基础神经网络的参数为共享参数，各个分支神经网络都可取用；此处对基础神经网络的类型以及结构不做限定。

在一个实施例中，如图 5 所示，上述步骤 43，根据样本图像的图像类别以及对应的实际图像类别，确定待训练的图像识别模型的损失值，包括：

步骤 51，根据样本图像集中样本图像的图像类别以及对应的实际图像类别，确定样本图像集中样本图像的损失值；

步骤 52，根据多个分支神经网络确定出的样本图像集中样本图像的损失值，得到与样本图像集对应的损失值，并作为目标分类损失值；

步骤 53，获取与多个分支神经网络对应的样本图像子集中所有样本图像的损失值，将样本图像子集中所有样本图像的损失值的和作为与多个分支神经网络对应的分类损失值；

步骤 54，根据目标分类损失值以及与多个分支神经网络对应的分类损失值，计算得到待训练的图像识别模型的损失值。

具体地，以样本图像集包括三个样本图像子集 **head classes**，**medium classes**，**tail classes** 为例进行说明；目标分类损失值是将整个待训练的图像识别模型输出的图像类别（即多个分支神经网络输出的图像类别的融合结果）与样本图像集中样本图像的实际类别进行比较后得到的损失值；因此，目标分类损失值计算的是三个分支神经网络输出的与样

本图像集中样本图像对应的图像类别；再将全部图像类别与实际图像类别输入损失函数，得到的损失值即为目标分类损失值，如下式所示：

$$L_f = J(F_{net}(X), Y), \text{ where } F_{net}(X) = N_{\bar{h}+\bar{m}+\bar{t}}(X) + N_{\bar{m}+\bar{t}}(X) + N_{\bar{t}}(X);$$

其中， $L_f$  为目标分类损失值； $J$  为交叉熵损失函数， $F_{net}$  为待训练的图像识别模型； $X$  为样本图像输入序列中的样本图像， $Y$  为样本图像的实际图像类别； $h$ 、 $m$ 、 $t$  分别为图像数量依次递减的第一、第二以及第三样本图像子集； $N_{\bar{h}+\bar{m}+\bar{t}}$ 、 $N_{\bar{m}+\bar{t}}$ 、 $N_{\bar{t}}$  为与三个样本图像子集对应的三个分支神经网络，角标为分支神经网络对应的样本图像子集。

分类损失值是各个分支神经网络针对其对应的样本图像子集得到的损失值，并不是针对整个样本图像集。例如  $N_{\bar{h}+\bar{m}+\bar{t}}$  分支神经网络与第一、第二以及第三样本图像子集为对应关系，那么计算  $N_{\bar{h}+\bar{m}+\bar{t}}$  分支神经网络时相当于计算其与整个样本图像集的损失值。而  $N_{\bar{t}}$  只与第三样本图像子集存在对应关系，因此计算  $N_{\bar{t}}$  的损失值时只需根据第三样本图像子集中相应样本图像的实际图像类别计算损失值即可；得到全部分支神经网络计算出的分类损失值后进行加和运算，结果即为最终的图像类别的预测结果；具体如下式：

$$L_i = \sum_{x_i \in X} J(N_{\bar{h}+\bar{m}+\bar{t}}(x_i), y_i) + \sum_{x_i \in S_{m+t}} J(N_{\bar{m}+\bar{t}}(x_i), y_i) + \sum_{x_i \in S_t} J(N_{\bar{t}}(x_i), y_i);$$

其中， $L_i$  为多个分支神经网络对应的分类损失值的和； $s_{m+t}$  为  $X$  的一个子集，包含样本图像输入序列中属于第二以及第三样本图像子集的样本图像； $s_t$  为  $X$  的另一个子集，包含样本图像输入序列中属于第三样本图像子集的样本图像。

待训练的图像识别模型的损失值由分类损失值和目标分类损失值共同计算得到，具体如下式：

$$L_{all} = (1-\alpha)L_f / n_1 + \alpha L_i / n_2;$$

其中， $L_{all}$  为待训练的图像识别模型的损失值； $\alpha$  为超参数； $n_1$  是  $X$  中样本图像的数量； $n_2$  是  $X$ 、 $S_{m+t}$  以及  $S_t$  中样本图像数量的总和。

需要说明的是，对于长尾程度不同的数据集，可以通过  $L_{all}$  函数中的超参数  $\alpha$  来调整；另外，当数据集为正常分布状态时（即各个图像类别的图像数量较为平均），将超参数  $\alpha$  设置为 0 即可正常运行。

上述实施例通过样本图像的图像类别以及对应的实际图像类别的差异，计算出目标分

类损失值以及分类损失值，并进一步得到待训练的图像识别模型的损失值，以此能够对待训练的图像识别模型中的参数进行调整，使得训练过程中图像数量较少的图像类别也能得到充分训练，避免了传统神经网络训练过程中对于长尾数据的忽视，提升了图像识别模型生成的效果。

在一个实施例中，如图 6 所示，上述步骤 21，在获取样本图像集之前，还包括：

步骤 61，获取样本图像，根据样本图像的图像类别，确定图像类别的图像数量；

步骤 62，根据图像类别的图像数量，得到图像类别的排列顺序，按照排列顺序将图像类别分为多个类别组合；多个类别组合包含相同数量的图像类别；

步骤 63，根据多个类别组合，以及与多个类别组合中的图像类别对应的样本图像，得到与多个类别组合对应的样本图像子集；将多个样本图像子集的组合作为样本图像集。

具体地，首先服务器从终端处获取样本图像，识别样本图像的图像类别，按照图像类别将样本图像进行归类，并统计各个图像类别对应样本图像的数量。根据图像类别对应样本图像的数量，对图像类别进行排序，从高到低依次排列，得到排列顺序。根据分支神经网络的数量以及图像类别的数量，将图像类别平均分配至多个类别组合。例如 3 个分支神经网络，6 个图像类别，则可两个图像类别分为一组，得到三个类别组合。根据类别组合以及类别组合对应的样本图像得到与类别组合对应的样本图像子集，多个样本图像子集即组合为样本图像集。

本实施例根据图像类别的图像数量，按照从高到低或从低到高的顺序进行排列，根据排列顺序平均分组，得到包含有相同数量图像类别的样本图像子集，实现了样本图像的预处理，各个分支神经网络按照长尾数据分布的特性与样本图像子集相互对应；使得训练过程中图像数量较少的图像类别也能得到充分训练，避免了传统神经网络训练过程中对于长尾数据的忽视，提升了图像识别模型生成的效果。

应该理解的是，虽然图 2、4-6 的流程图中的各个步骤按照箭头的指示依次显示，但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明，这些步骤的执行并没有严格的顺序限制，这些步骤可以以其它的顺序执行。而且，图 2、4-6 中的至少一部分步骤可以包括多个步骤或者多个阶段，这些步骤或者阶段并不必然是在同一时刻执行完成，而是可以在不同的时刻执行，这些步骤或者阶段的执行顺序也不必然是依次进行，而是可以与其它步骤或者其它步骤中的步骤或者阶段的至少一部分轮流或者交替地执行。

在一个实施例中，如图 7 所示，提供了一种图像识别模型生成装置，包括：

获取模块 71，用于获取样本图像集；样本图像集包括图像数量依次递减的多个样本图像子集，多个样本图像子集均包含相同数量的图像类别；

训练模块 72，用于根据样本图像集，对待训练的图像识别模型进行训练，得到待训练的图像识别模型的损失值；待训练的图像识别模型包括多个分支神经网络，每个分支神经网络用于对相应的图像进行识别；损失值包括目标分类损失值以及与各个分支神经网络对应的分类损失值，目标分类损失值为待训练的图像识别模型针对样本图像集的损失值，分类损失值为相应的分支神经网络针对与分支神经网络对应的样本图像子集的损失值；

调整模块 73，用于根据损失值调整待训练的图像识别模型的模型参数，直至损失值低于预设阈值时，将待训练的图像识别模型作为训练完成的图像识别模型。

在一个实施例中，训练模块 72 还用于对样本图像集中的多个样本图像子集进行均匀采样，得到样本图像输入序列；根据样本图像输入序列，将样本图像输入待训练的图像识别模型，得到样本图像的图像类别；根据样本图像的图像类别以及对应的实际图像类别，确定待训练的图像识别模型的损失值。

在一个实施例中，训练模块 72 还用于将样本图像输入待训练的图像识别模型，使得基础神经网络获取样本图像的第一图像特征，以及使得分支神经网络根据第一图像特征得到样本图像的第二图像特征，并根据第二图像特征确定样本图像集中样本图像的图像类别。

在一个实施例中，训练模块 72 还用于根据样本图像集中样本图像的图像类别以及对应的实际图像类别，确定样本图像集中样本图像的损失值；根据多个分支神经网络确定出的样本图像集中样本图像的损失值，得到与样本图像集对应的损失值，并作为目标分类损失值；获取与多个分支神经网络对应的样本图像子集中所有样本图像的损失值，将样本图像子集中所有样本图像的损失值的和作为与多个分支神经网络对应的分类损失值；根据目标分类损失值以及与多个分支神经网络对应的分类损失值，计算得到待训练的图像识别模型的损失值。

在一个实施例中，获取模块 71 还用于获取样本图像，根据样本图像的图像类别，确定图像类别的图像数量；根据图像类别的图像数量，得到图像类别的排列顺序，按照排列顺序将图像类别分为多个类别组合；多个类别组合包含相同数量的图像类别；根据多个类别组合，以及与多个类别组合中的图像类别对应的样本图像，得到与多个类别组合对应的样本图像子集；将多个样本图像子集的组合作为样本图像集。

关于图像识别模型生成装置的具体限定可以参见上文中对于图像识别模型生成方法的限定，在此不再赘述。上述图像识别模型生成装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。

在一个实施例中，提供了一种计算机设备，该计算机设备可以是服务器，其内部结构图可以如图 8 所示。该计算机设备包括通过系统总线连接的处理器、存储器和网络接口。其中，该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机程序和数据库。该内存储器为非易失性存储介质中的操作系统和计算机程序的运行提供环境。该计算机设备的数据库用于存储图像识别模型生成数据。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机程序被处理器执行时以实现一种图像识别模型生成方法。

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。

在一个实施例中，提供了一种计算机设备，包括存储器和处理器，存储器中存储有计算机程序，该处理器执行计算机程序时实现上述各方法实施例中的步骤。

在一个实施例中，提供了一种计算机可读存储介质，其上存储有计算机程序，计算机程序被处理器执行时实现上述各方法实施例中的步骤。

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，上述的计算机程序可存储于一非易失性计算机可读存储介质中，该计算机程序在执行时，可包括如上述各方法的实施例的流程。其中，本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和易失性存储器中的至少一种。非易失性存储器可包括只读存储器（Read-Only Memory，ROM）、磁带、软盘、闪存或光存储器等。易失性存储器可包括随机存取存储器（Random Access Memory，RAM）或外部高速缓冲存储器。作为说明而非局限，RAM 可以是多种形式，比如静态随机存取存储器（Static Random Access Memory，SRAM）或动态随机存取存储器（Dynamic Random Access Memory，DRAM）等。

以上实施例的各技术特征可以进行任意的组合，为使描述简洁，未对上述实施例中的各个技术特征所有可能的组合都进行描述，然而，只要这些技术特征的组合不存在矛盾，

都应当认为是本说明书记载的范围。

以上实施例仅表达了本申请的几种实施方式，其描述较为具体和详细，但并不能因此而理解为对发明专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本申请构思的前提下，还可以做出若干变形和改进，这些都属于本申请的保护范围。因此，本申请专利的保护范围应以所附权利要求为准。