

神经网络模型训练方法、装置、计算机设备及存储介质

本申请以 2019 年 1 月 4 日提交的申请号为 201910008317.2，名称为“神经网络模型训练方法、装置、计算机设备及存储介质”的中国发明专利申请为基础，并要求其优先权。

技术领域

本申请涉及神经网络领域，尤其涉及一种神经网络模型训练方法、装置、计算机设备及存储介质。

背景技术

目前深度学习算法在计算机视觉应用开发中占据重要地位，而深度学习算法对于训练数据有一定的要求，在训练数据量不足时，对于低频次困难样本（hard example）拟合效果欠佳。基于上述情况，传统上，有人提出了一些困难样本挖掘的训练方式，保留训练集中低频次、欠拟合的样本，去除高频次、易识别样本，从而达到精简训练集的目的，用于提高训练针对性，但是，发明人意识到，上述传统的方案中，一方面是减少了训练集中的训练数据，不利于模型的训练，另一方面是即使对训练数据进行增益或者补充，也难做到模型训练中训练数据针对性的增强，无法直接分析模型训练过程中所欠缺的样本，也就是困难样本，从而导致上述传统的训练方式的针对性和训练效率都比较低。

发明内容

本申请提供了一种神经网络模型训练方法、装置、计算机设备及存储介质，选出了具备针对性的训练样本，且提高了模型训练的针对性以及训练效率。

一种神经网络模型训练方法，包括：根据训练集的训练样本对深度神经网络模型进行训练，以获得训练后的深度神经网络模型；根据所述训练后的深度神经网络模型对参考集合的所有参考样本进行数据验证，以获得所述所有参考样本中每个参考样本的模型预测值，所述参考集合包括验证集和/或测试集；计算所述每个参考样本的模型预测值与所述每个参考样本对应的真实标注之间的差异衡量指标，所述每个参考样本预先进行了数据标注；将所有所述参考样本中差异衡量指标低于或等于预设阈值的目标参考样本作为比较样本；计算所述训练集中的训练样本与每个所述比较样本之间的相似度；将与所述比较样本之间的相似度满足预设扩增条件的训练样本作为待扩增样本；对所述待扩增样本进行数据扩增以获得目标训练样本；将所述目标训练样本作为所述训练集中的训练样本对所述训练后的深度神经网络模型进行训练，直至所述验证集所有的验证样本的模型预测值满足预设

训练结束条件。

一种神经网络模型训练装置，包括：训练模块，用于根据训练集的训练样本对深度神经网络模型进行训练，以获得训练后的深度神经网络模型；验证模块，用于根据所述训练模块训练得到的所述训练后的深度神经网络模型对参考集合的所有参考样本进行数据验证，以获得所述所有参考样本中每个参考样本的模型预测值，所述参考集合包括验证集和/或测试集；第一计算模块，用于计算所述每个参考样本的模型预测值与所述每个参考样本对应的真实标注之间的差异衡量指标，所述每个参考样本预先进行了数据标注；第一确定模块，用于将所有所述参考样本中所述第一计算模块计算得到的差异衡量指标低于或等于预设阈值的目标参考样本作为比较样本；第二计算模块，用于计算所述训练集中的训练样本与所述第一确定模块确定的每个所述比较样本之间的相似度；第二确定模块，用于将与所述第二计算模块计算得到的所述比较样本之间的相似度满足预设扩增条件的训练样本作为待扩增样本；扩增模块，用于对所述第二确定模块确定的所述待扩增样本进行数据扩增以获得目标训练样本；所述训练模块，用于将所述扩增样本扩增得到的所述目标训练样本作为所述训练集中的训练样本对所述训练后的深度神经网络模型进行再次训练，直至所述验证集所有的验证样本的模型预测值满足预设训练结束条件。

一种计算机设备，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可读指令，所述处理器执行所述计算机可读指令时实现上述神经网络模型训练方法所对应的步骤。

一个或多个存储有计算机可读指令的非易失性可读存储介质，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器执行上述神经网络模型训练方法所对应的步骤。

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，本申请的其他特征和优点将从说明书、附图以及权利要求变得明显。

附图说明

为了更清楚地说明本申请实施例的技术方案，下面将对本申请实施例的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实施例的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图1是本申请中神经网络模型训练方法的架构示意图；

图2是本申请中神经网络模型训练方法的实施例流程示意图；

图 3 是本申请中神经网络模型训练方法的实施例流程示意图；

图 4 是本申请中神经网络模型训练方法的实施例流程示意图；

图 5 是本申请中神经网络模型训练方法的实施例流程示意图；

图 6 是本申请中神经网络模型训练方法的实施例流程示意图；

图 7 是本申请中神经网络模型训练装置的一实施例结构示意图；

图 8 是本申请中计算机设备的一结构示意图。

具体实施方式

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，所描述的实施例是本申请实施例一部分实施例，而不是全部的实施例。基于本申请实施例中的实施例，本领域普通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，都属于本申请实施例保护的范围。

本申请提供了一种神经网络模型训练方法，可应用在如图 1 的架构示意图中，神经网络模型训练装置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，或者该神经网络模型训练装置作为独立的装置，或集成在上述服务器中实现，这里不做限定。服务器可以获取用于进行模型训练的训练集中的训练样本和参考样本，根据训练集的训练样本对深度神经网络模型进行训练，以获得训练后的深度神经网络模型；根据所述训练后的深度神经网络模型对参考集合的所有参考样本进行数据验证，以获得所述所有参考样本中每个参考样本的模型预测值，所述参考集合包括验证集和/或测试集；计算所述每个参考样本的模型预测值与所述每个参考样本对应的真实标注之间的差异衡量指标；将所有所述参考样本中差异衡量指标低于或等于预设阈值的目标参考样本作为比较样本；计算所述训练集中的训练样本与每个所述比较样本之间的相似度；将与所述比较样本之间的相似度满足预设扩增条件的训练样本作为待扩增样本；对所述待扩增样本进行数据扩增以获得目标训练样本；将所述目标训练样本作为所述训练集中的训练样本对所述训练后的深度神经网络模型进行训练，直至所述验证集所有的验证样本的模型预测值满足预设训练结束条件。由以上方案可看出，由于是针对性的挑选被扩增的样本数据，使得扩增模型训练的训练样本数据，并且是将测试集和/或验证集中的样本的预测结果参与到模型训练中，与验证集、测试集产生直接交互，从结果上去直接分析模型训练过程中所欠缺的样本，也就是困难样本，使得选出了具备针对性的训练样本，从而提高了模型训练的针对性以及训练效率。下面对本申请进行详细的描述：

请参阅图 2，图 2 为本申请中一种深度神经网络模型训练方法一个实施例流程示意图，

包括如下步骤:

S10: 根据训练集的训练样本对深度神经网络模型进行训练, 以获得训练后的深度神经网络模型。

训练集是深度神经网络模型训练的基础, 深度神经网络模型可以想象为一个强大的非线性拟合器, 去对训练集上的数据, 也即训练样本进行拟合。因此, 在预备好了训练集后, 可以根据训练集的训练样本对深度神经网络模型进行训练, 以获得训练后的深度神经网络模型。其中, 需要说明的是, 上述深度神经网络模型指的是卷积神经网络模型, 也可以是循环神经网络模型, 还可以是其他类型的卷积神经网络模型, 本申请实施例不做限定。另外, 上述训练过程为有效监督训练过程, 训练集中的训练样本是已经进行预设的标注。示例性的, 若是为了训练出用于图片分类的深度神经网络模型, 则会对训练样本进行图片分类标注, 从而训练出用于图片分类的深度神经网络模型, 例如用于对病灶图像分类的深度神经网络模型。

具体地, 本申请实施例可以预设训练时期(epoch), 示例性的, 可以将 10 个 epoch 作为一次完整的训练周期, 其中, 每一个 epoch 指的是, 根据训练集的所有训练样本对深度神经网络模型进行训练一次, 每个 10 个 epoch 指的是根据训练集的所有训练样本对深度神经网络模型进行训练 10 次。需要说明的是, 具体的 epoch 个数本申请实施例不做限定, 示例性的, 还可以将 8 个时期作为一次完整的训练周期。

S20: 根据所述训练后的深度神经网络模型对参考集合的所有参考样本进行数据验证, 以获得所述所有参考样本中每个参考样本的模型预测值, 所述参考集合包括验证集和/或测试集。

验证集指的是: 本申请实施例中整个训练过程中对深度神经网络模型的有效性进行评估的样本数据。当深度神经网络模型训练进行到一定程度的时候就会使用验证集上的样本数据去校验深度神经网络模型, 以防深度神经网络模型发生过拟合, 所以验证集上的样本数据间接的参与到了模型训练过程中, 从而根据验证结果确定深度神经网络模型此刻的训练状态是否对训练训练集以外的数据有效。而测试集是最终用于评深度神经网络模型准确率的样本数据。

在本申请实施例中, 将上述验证集和/或测试集作为参考集合, 将验证集和/或测试集的样本数据作为参考集合中的参考样本。示例性的, 在训练每 10 个 epoch 结束后, 可以得到训练后的深度神经网络模型, 此时, 根据所述训练后的深度神经网络模型对参考集合的所有参考样本进行数据验证, 以获得所述所有参考样本中每个参考样本的模型预测值。需

要说明的是，模型预测值指的是，经过一定的训练后，用于神经网络模型对参考样本进行验证所产生的验证结果，示例性的，若该神经网络模型用于图像分类，则该模型预测值用于表征图像分类的准确度。

S30：计算所述每个参考样本的模型预测值与所述每个参考样本对应的真实标注之间的差异衡量指标，所述每个参考样本预先进行了数据标注。

在获得所述所有参考样本中每个参考样本的模型预测值后，计算所有参考样本中，每个参考样本的模型预测值与参考样本对应的真实标注之间的差异衡量指标。

可以理解，作为一种有效监督训练方式，验证集或测试集中的样本数据都预先进行了数据标注，也即每个参考样本对应的真实标注，差异衡量指标是用于表征参考样本的模型预测值与该参考样本对应的真实标注之间差异程度的指标。示例性的，对于参考样本 A，神经网络模型预测出来的模型预测值是[0.8,5,0,0.2,0,0]，而真实标注是[1,0,0,0,0]，则可以根据这两组数据进行计算，得到差异衡量指标，这样就可以知道模型预测值与真实标注究竟有多少差距。

在一实施方式中，如图 3 所示，步骤 S30 中，也即所述计算所述每个参考样本的模型预测值与所述每个参考样本对应的真实标注之间的差异衡量指标，包括如下步骤：

S31：确定所述训练后的神经网络模型所采用的差异衡量指标类型。

应当理解，在根据所述差异衡量指标类型，计算所述每个参考样本的模型预测值与所述每个参考样本对应的真实标注之间的差异衡量指标之前，本方案需先确定所述训练后的神经网络模型所采用的差异衡量指标类型，具体取决于训练后的神经网络模型的作用，神经网络模型的作用指的是该神经网络模型是用于图像分割或图像分类等作用，依据不同的神经网络模型的作用选择合适的差异衡量指标类型。

在一实施例中，如图 4 所示，步骤 S31 中，也即所述确定所述训练后的神经网络模型所采用的差异衡量指标类型，包括如下步骤：

S311：获取预设指标对应列表，所述预设指标列表包含差异衡量指标类型与模型作用指示字符之间的对应关系，所述模型作用指示字符用于指示神经网络模型的作用。

所述模型作用指示字符可以指示神经网络模型的作用，具体可以用数字、字母等方式自定义，这里不做限定。具体地，所述差异衡量指标类型包括交叉熵系数、杰卡德系数以及 dice 系数，其中，指示神经网络模型用于图像分类作用的模型作用指示字符与所述交叉熵系数相对应，指示神经网络模型用于图像分割作用的模型作用指示字符与所述杰卡德系数或 dice 系数相对应。

S312: 确定所述训练后的神经网络模型对应的模型作用指示字符。

S313: 根据所述差异衡量指标与模型作用指示字符之间的对应关系, 以及所述训练后的神经网络模型对应的模型作用指示字符, 确定所述训练后的神经网络模型所采用的差异衡量指标类型。

对于步骤 S312-S313, 可以理解, 获取预设指标对应列表之后, 可以根据预设指标对应列表确定出根据所述差异衡量指标与模型作用指示字符之间的对应关系, 因此, 可以根据所述训练后的神经网络模型对应的模型作用指示字符确定出所述训练后的神经网络模型所采用的差异衡量指标类型。

S32: 根据所述差异衡量指标类型, 计算所述每个参考样本的模型预测值与所述每个参考样本对应的真实标注之间的差异衡量指标。

举例说明, 假设本申请实施例中的神经网络模型对应的模型作用为用于图像分类, 则可以将交叉熵系数作为每个参考样本的模型预测值与所述每个参考样本对应的真实标注之间的差异衡量指标。

假设现在有一个参考样本的真实标注的分布为 $p(x)$, 参考样本的模型预测值为 $q(x)$, 也即训练后的神经网络模型的预测分布为 $q(x)$, 则可以依据以下公式计算真实标注和模型预测值之间的交叉熵 $H(p, q)$:

$$H(p, q) = \sum_x p(x) \cdot \log\left(\frac{1}{q(x)}\right);$$

需要说明的是, 假设本申请实施例中的神经网络模型对应的模型作用为用于图像分割, 则可以计算真实标注和模型预测值之间的根据杰卡德系数或 dice 系数作为实标注和模型预测值之间的差异衡量指标, 具体计算过程这里不做详细描述。

S40: 将所有所述参考样本中差异衡量指标低于或等于预设阈值的目标参考样本作为比较样本。

可以理解, 经过步骤 S30 后, 可以得到参考集合所有参考样本中, 每个参考样本对应的差异衡量指标, 在本申请实施例中, 将所述所有参考样本中差异衡量指标低于或等于预设阈值的目标参考样本作为比较样本, 用于后续参与训练样本的相似度计算。可以理解, 此时得到的比较样本就是上述所提到的苦难样本, 并且得到的比较样本可以为一个或多个, 具体由神经网络模型的训练情况确定。需要说明的是, 预设阈值是根据项目要求或实际经验来定, 具体这里不做限定, 示例性的, 以神经网络模型为用于图像分割的

模型为例，上述预设阈值可设定为 0.7。

S50：计算所述训练集中的训练样本与每个所述比较样本之间的相似度。

在得到比较样本后，计算所述训练集中的训练样本与每个所述比较样本之间的相似度。为了便于理解，这里举个简单的例子进行说明，示例性的，假设比较样本有 3 个，训练样本有 10 个，则可以分别计算出每个比较样本与 10 个训练样本中每个训练样本的相似度，共 30 个相似度。

在一实施例中，如图 5 所示，步骤 S50 中，也即所述计算所述训练集中的训练样本与所述比较样本之间的相似度，包括如下步骤：

S51：根据预设特征提取模型对所述训练集的每个训练样本进行特征提取以获得每个训练样本的特征向量，所述预设图像特征提取模型为基于卷积神经网络所训练得到的特征提取模型。

S52：根据所述预设特征提取模型对所述比较样本进行特征提取以获得每个比较样本的特征向量。

S53：根据所述每个训练样本的特征向量与所述每个比较样本的特征向量计算所述训练集中的训练样本与所述比较样本之间的相似度。

对于步骤 S51-S53，本申请实施例基于特征向量的方式计算计算所述训练集中的训练样本与所述比较样本之间的相似度。其中，基于卷积神经的图像特征向量提取，不同的图像相似算法最终找到的图片有效性有所不同，就有较高的针对性，有利于模型的训练。

在一实施例中，如图 6 所示，步骤 S53，也即步骤所述根据所述每个训练样本的特征向量与所述每个比较样本的特征向量计算所述训练集中的训练样本与所述比较样本之间的相似度，包括如下步骤：

S531：计算所述每个训练样本的特征向量与所述每个比较样本的特征向量之间的余弦距离。

S532：将所述每个训练样本的特征向量与所述每个比较样本的特征向量之间的余弦距离作为所述每个训练样本与所述每个比较样本之间的相似度。

对于步骤 S531-S532，可以理解，除了上述以余弦距离来表征训练样本与比较样本之间的相似度外，还可以计算每个训练样本的特征向量与所述每个比较样本的特征向量得到的欧式距离、曼哈顿距离等用于表征上述相似度，具体本申请实施例不做限定。这里，以余弦相似度计算方式为例，假设训练样本对应的特征向量为 $x_i, i \in (1, 2, \dots, n)$ ，比较样本对应

的特征向量为 $y_i, i \in (1, 2, \dots, n)$, 其中, n 为正整数, 则训练样本的特征向量与所述每个比

$$\cos(\theta) = \frac{\sum_{i=1}^n (x_i \times y_i)}{\sqrt{\sum_{i=1}^n (x_i)^2} \times \sqrt{\sum_{i=1}^n (y_i)^2}}。$$

较样本的特征向量之间的余弦距离为:

S60: 将与所述比较样本之间的相似度满足预设扩增条件的训练样本作为待扩增样本。

在计算所述训练集中的训练样本与每个所述比较样本之间的相似度之后, 将与所述比较样本之间的相似度满足预设扩增条件的训练样本作为待扩增样本。其中, 需要说明的是, 上述预设扩增条件可以依据实际应用场景进行调整。示例性的, 若训练集中的训练样本与所述比较样本之间的相似度排在前 3 位, 则排前 3 位的训练样本满足上述预设扩增条件。举例说明, 例如存在比较样本 1 和比较样本 2, 计算比较样本 1 与训练集中每个训练样本的相似度, 将相似度排在前 3 位的训练样本作为待扩增样本; 同理计算比较样本 2 与训练集中每个训练样本的相似度, 将相似度排在前 3 位的训练样本作为待扩增样本, 其他比较样本确定出待扩增样本的方式类似, 从而可以得到每个比较样本确定出的待扩增样本。可以理解, 上述得到的待扩增样本为与比较样本最为相似的一组样本。

可以看出, 这里根据不同的应用场景, 可寻找全局最高相似度、局部最高相似度以契合需求, 整个过程无需人为观测、人为遴选样本, 是一种高效的筛选机制。

S70: 对所述待扩增样本进行数据扩增以获得目标训练样本。

在得到与所述比较样本之间的相似度满足预设扩增条件的训练样本作为待扩增样本后, 对所述待扩增样本进行数据扩增以获得目标训练样本。需要说明的是, 本申请实施例可以采用常规的图像扩增方式对被确定出来的待扩增样本进行统一的数据扩增, 示例性的, 可以以两倍数据增强 (例如旋转、平移、放缩等) 等方式进行扩增, 扩增后的样本, 也就是目标训练样本。这里可以减少数据增益总量, 仅增益少部分数据, 便于提升模型训练效率。

S80: 将所述目标训练样本作为所述训练集中的训练样本对所述训练后的深度神经网络模型进行训练, 直至所述验证集所有的验证样本的模型预测值满足预设训练结束条件。

在得到扩增后的样本后, 也即目标训练样本后, 将所述目标训练样本作为所述训练集中的训练样本对所述训练后的深度神经网络模型进行训练, 直至所述验证集所有的验证样本的模型预测值满足预设训练结束条件。也就是说, 在得到扩增得到的目标训练样本后, 再次将目标训练样本作为训练集以验证集的样本数据对深度神经网络模型进行训练, 周而复始, 开始新一轮训练, 基于此种操作, 实现从模型预测的结果出发, 返回源头进行优化

并达到改善预测结果的目的，从而提高模型预测性能，提高了模型训练效率。

在一实施方式中，将上述目标训练样本按照一定的比例分配至训练集合验证集中，示例性的，使得上述分配结果为训练集中的样本与验证集中的样本比例保持在 5:1 左右，抑或是其他分配比例，这里不做限定。

在一实施方式中，所述将所述目标训练样本作为所述训练集中的训练样本对所述训练后的深度神经网络模型进行训练，直至所述验证集所有的验证样本的模型预测值满足预设训练结束条件，包括：将所述目标训练样本作为所述训练集中的训练样本对所述训练后的深度神经网络模型进行训练，直至所述验证集所有的验证样本的每个验证样本的对应的差异衡量指标低于或等于所述预设阈值。除此之外，还可以有其他的预设训练结束条件，例如模型的训练迭代的次数已经达到了预设上限，具体这里也不做限定。

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。

在一实施例中，提供一种神经网络模型训练装置，该神经网络模型训练装置与上述实施例中神经网络模型训练方法一一对应。如图 7 所示，该神经网络模型训练装置 10 包括训练模块 101、验证模块 102、第一计算模块 103、第一确定模块 104、第二计算模块 105、第二确定模块 106、扩增模块 107、各功能模块详细说明如下：训练模块 101，用于根据训练集的训练样本对深度神经网络模型进行训练，以获得训练后的深度神经网络模型；验证模块 102，用于根据所述训练模块 101 训练得到的所述训练后的深度神经网络模型对参考集合的所有参考样本进行数据验证，以获得所述所有参考样本中每个参考样本的模型预测值，所述参考集合包括验证集和/或测试集；第一计算模块 103，用于计算验证模块 102 验证得到的所述每个参考样本的模型预测值与所述每个参考样本对应的真实标注之间的差异衡量指标，所述每个参考样本预先进行了数据标注；第一确定模块 104，用于将所有所述参考样本中所述第一计算模块 103 计算得到的差异衡量指标低于或等于预设阈值的目标参考样本作为比较样本；第二计算模块 105，用于计算所述训练集中的训练样本与所述第一确定模块 104 确定的每个所述比较样本之间的相似度；第二确定模块 106，用于将与所述第二计算模块 105 计算得到的所述比较样本之间的相似度满足预设扩增条件的训练样本作为待扩增样本；扩增模块 107，用于对所述第二确定模块 106 确定的所述待扩增样本进行数据扩增以获得目标训练样本；所述训练模块 101，用于将所述扩增样本扩增得到的所述目标训练样本作为所述训练集中的训练样本对所述训练后的深度神经网络模型进行再次训练，直至所述验证集所有的验证样本的模型预测值满足预设训练结束条件。

在一实施例中，所述训练模块 101 用于所述将所述目标训练样本作为所述训练集中的训练样本对所述训练后的深度神经网络模型进行训练，直至所述验证集所有的验证样本的模型预测值满足预设训练结束条件，具体包括：所述训练模块 101 用于：将所述目标训练样本作为所述训练集中的训练样本对所述训练后的深度神经网络模型进行训练，直至所述验证集所有的验证样本的每个验证样本的对应的差异衡量指标低于或等于所述预设阈值。

在一实施例中，第一计算模块 103 具体用于：确定所述训练后的深度神经网络模型所采用的差异衡量指标类型；根据所述差异衡量指标类型，计算所述每个参考样本的模型预测值与所述每个参考样本对应的真实标注之间的差异衡量指标。

在一实施例中，第一计算模块 103 用于确定所述训练后的深度神经网络模型所采用的差异衡量指标类型，具体包括：第一计算模块 103 具体用于：获取预设指标对应列表，所述预设指标列表包含差异衡量指标类型与模型作用指示字符之间的对应关系，所述模型作用指示字符用于指示深度神经网络模型的作用；确定所述训练后的深度神经网络模型对应的模型作用指示字符；根据所述差异衡量指标与模型作用指示字符之间的对应关系，以及所述训练后的深度神经网络模型对应的模型作用指示字符，确定所述训练后的深度神经网络模型所采用的差异衡量指标类型。

在一实施例中，所述差异衡量指标类型包括交叉熵系数、杰卡德系数以及 dice 系数，其中，指示深度神经网络模型用于图像分类作用的模型作用指示字符与所述交叉熵系数相对应，指示深度神经网络模型用于图像分割作用的模型作用指示字符与所述杰卡德系数或 dice 系数相对应。

在一实施例中，第二计算模块 105，具体用于：根据预设特征提取模型对所述训练集的每个训练样本进行特征提取以获得每个训练样本的特征向量，所述预设图像特征提取模型为基于卷积神经网络所训练得到的特征提取模型；根据所述预设特征提取模型对所述比较样本进行特征提取以获得每个比较样本的特征向量；根据所述每个训练样本的特征向量与所述每个比较样本的特征向量计算所述训练集中的训练样本与所述比较样本之间的相似度。

在一实施例中，第二计算模块 105 用于根据所述每个训练样本的特征向量与所述每个比较样本的特征向量计算所述训练集中的训练样本与所述比较样本之间的相似度，包括：

第二计算模块 105 用于：计算所述每个训练样本的特征向量与所述每个比较样本的特征向量之间的余弦距离；将所述每个训练样本的特征向量与所述每个比较样本的特征向量之间的余弦距离作为所述每个训练样本与所述每个比较样本之间的相似度。

关于神经网络训练装置装置的具体限定可以参见上文中对于神经网络训练装置方法的限定，在此不再赘述。上述神经网络训练装置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、网络接口和数据库。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临时存储训练样本、参考样本等。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神经网络训练方法。

在一个实施例中，提供了一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机可读指令，处理器执行计算机可读指令时实现以下步骤：根据训练集的训练样本对深度神经网络模型进行训练，以获得训练后的深度神经网络模型；根据所述训练后的深度神经网络模型对参考集合的所有参考样本进行数据验证，以获得所述所有参考样本中每个参考样本的模型预测值，所述参考集合包括验证集和/或测试集；计算所述每个参考样本的模型预测值与所述每个参考样本对应的真实标注之间的差异衡量指标，所述每个参考样本预先进行了数据标注；将所有所述参考样本中差异衡量指标低于或等于预设阈值的目标参考样本作为比较样本；计算所述训练集中的训练样本与每个所述比较样本之间的相似度；将与所述比较样本之间的相似度满足预设扩增条件的训练样本作为待扩增样本；对所述待扩增样本进行数据扩增以获得目标训练样本；将所述目标训练样本作为所述训练集中的训练样本对所述训练后的深度神经网络模型进行训练，直至所述验证集所有的验证样本的模型预测值满足预设训练结束条件。

在一个实施例中，提供了一个或多个存储有计算机可读指令的非易失性可读存储介质，该非易失性可读存储介质上存储有计算机可读指令，该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实现以下步骤：根据训练集的训练样本对深度神经网络模型进行训练，以获得训练后的深度神经网络模型；根据所述训练后的深度神经网络模型对参考集合的所有参考样本进行数据验证，以获得所述所有参考样本中每个参考样

本的模型预测值,所述参考集合包括验证集和/或测试集;计算所述每个参考样本的模型预测值与所述每个参考样本对应的真实标注之间的差异衡量指标,所述每个参考样本预先进行了数据标注;将所有所述参考样本中差异衡量指标低于或等于预设阈值的目标参考样本作为比较样本;计算所述训练集中的训练样本与每个所述比较样本之间的相似度;将与所述比较样本之间的相似度满足预设扩增条件的训练样本作为待扩增样本;对所述待扩增样本进行数据扩增以获得目标训练样本;将所述目标训练样本作为所述训练集中的训练样本对所述训练后的深度神经网络模型进行训练,直至所述验证集所有的验证样本的模型预测值满足预设训练结束条件。

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,所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/或易失性存储器。非易失性存储器可包括只读存储器(ROM)、可编程ROM(PROM)、电可编程ROM(EPROM)、电可擦除可编程ROM(EEPROM)或闪存。易失性存储器可包括随机存取存储器(RAM)或者外部高速缓冲存储器。作为说明而非局限,RAM以多种形式可得,诸如静态RAM(SRAM)、动态RAM(DRAM)、同步DRAM(SDRAM)、双数据率SDRAM(DDRSDRAM)、增强型SDRAM(ESDRAM)、同步链路(Synchlink)DRAM(SLDRAM)、存储器总线(Rambus)直接RAM(RDRAM)、直接存储器总线动态RAM(DRDRAM)、以及存储器总线动态RAM(RDRAM)等。

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。

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实施例进行了详细的说明,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实施例各实施例技术方案的精神和范围,均应包含在本申请实施例的保护范围之内。