

一种用于神经网络的训练方法以及相关设备

本申请要求于 2020 年 7 月 15 日提交中国专利局、申请号为 202010680648.3、发明名称为“一种用于神经网络的训练方法以及相关设备”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

本申请涉及人工智能领域，尤其涉及一种用于神经网络的训练方法以及相关设备。

背景技术

人工智能 (Artificial Intelligence, AI) 是利用计算机或者计算机控制的机器模拟、延伸和扩展人的智能。人工智能包括研究各种智能机器的设计原理与实现方法，使机器具有感知、推理与决策的功能。目前，对基于深度学习的神经网络进行小样本学习 (few-shot learning) 是人工智能一种可行的研究方向。

元学习 (meta-learning) 是小样本学习研究中的一个重要的分支。元学习的主要思想是当目标任务的训练样本较少时，通过使用大量与目标小样本任务相似的小样本任务来训练神经网络，以使训练后的神经网络在目标任务上有着不错的初始值，然后利用少量的目标小样本任务的训练样本对训练后的神经网络进行调整。

但最后得到的神经网络输出的预测结果的精度依旧较低，因此，一种在采用元学习的方式进行训练的前提下，能够提高神经网络输出的预测结果的精度的训练方案亟待推出。

发明内容

本申请实施例提供了一种用于神经网络的训练方法以及相关设备，在一个小样本任务的训练过程中不仅学习到一个小样本任务中支持样本的特征，而且学习到 M 组支持样本的特征，也即神经网络在一个小样本学习中不仅能够学习到一个小样本任务内的相对关系，而且能够学习到多个小样本任务之间的绝对关系，从而提高输出的预测结果的精度。

为解决上述技术问题，本申请实施例提供以下技术方案：

第一方面，本申请实施例提供一种神经网络的训练方法，可用于人工智能领域的元学习领域中。方法包括：训练设备从样本集合中获取第一样本子集合，样本集合包括 M 组支持样本和 M 组查询样本，第一样本子集合包括第一支持样本和第一查询样本；训练设备还获取每个第一支持样本的标注结果。训练设备将第一样本子集合和每个第一支持样本的标注结果输入至第一神经网络，第一神经网络生成第一查询样本的第一特征信息，并得到第一神经网络输出的与第一查询样本对应的第一预测结果，第一预测结果为根据第一支持样本的特征信息和第一特征信息之间的相似度以及每个第一支持样本的标注结果生成的。训练设备根据与 M 组支持样本对应的 M 个第二特征信息、与 M 组支持样本对应的 M 个第一标注结果和第一查询样本的第一特征信息，生成与第一查询样本对应的第二预测结果。 M 个第二特征信息中的每一个均指示 M 组支持样本中一组支持样本的特征， M 个第一标注结果中的每一个均指示与 M 组支持样本中一组支持样本对应的正确结果。训练设备根据第一

损失函数和第二损失函数，对第一神经网络进行迭代训练，直至满足预设条件。其中，第一损失函数指示第一预测结果和第二标注结果之间的相似度，第二损失函数指示第一预测结果与第二预测结果之间的相似度，或者，第二损失函数指示第二预测结果和第二标注结果之间的相似度，第二标注结果为第一查询样本的正确结果。第一损失函数和第二损失函数均可以为交叉熵损失函数或相对熵损失函数。

本实现方式中，由于最后得到的神经网络过多学习同一小样本任务中不同样本在特征维度的区别，而忽略了与多个不同的小样本任务对应的多个不同类别的样本在特征维度的区别，从而导致神经网络输出的预测结果的精度较低。为了解决前述问题，本方案的训练阶段中，在每次小样本任务中，还会根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的特征信息，生成与第一查询样本对应的第二预测结果，并增设第二损失函数指示第一预测结果与第二预测结果之间的相似度，或者，指示第二预测结果和正确结果之间的相似度。从而使神经网络在一个小样本任务的训练过程中不仅能够学习到一个小样本任务中支持样本的特征，而且能够学习到 M 组支持样本的特征，也即神经网络在一个小样本学习中不仅能够学习到一个小样本任务中不同样本在特征维度的区别，而且能够学习到与多个不同小样本任务对应的多个不同类别的样本在特征维度的区别，从而提高输出的预测结果的精度。

在第一方面的一种可能实现方式中，训练设备将第一样本子集合输入至第一神经网络之后，方法还包括：训练设备通过第一神经网络生成第一支持样本的特征信息，第一支持样本选自于第一组，第一组为 M 组中的一个组；将第一支持样本的特征信息添加至特征信息库的第一队列中。其中，特征信息库包括与 M 组支持样本一一对应的 M 个队列，第一队列为 M 个队列中与第一组对应的一个队列，用于存储与第一组的支持样本对应的特征信息，第一队列还用于存储与第一组的支持样本对应的标注结果。训练设备根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一特征信息，生成与第一查询样本对应的第二预测结果之前，方法还包括：训练设备根据特征信息库，获取与 M 组支持样本对应的 M 个第二特征信息和与 M 组支持样本对应的 M 个第一标注结果。

本实现方式中，训练设备利用每次训练过程中生成的特征信息，来生成 M 组中每组支持样本的特征，提高了训练过程的效率。

在第一方面的一种可能实现方式中，第二队列为 M 个队列中任一队列，第二队列包括至少一个支持样本的特征信息。训练设备根据特征信息库，获取与 M 组支持样本对应的第二特征信息，包括：训练设备对第二队列包括的所有支持样本的特征信息进行取平均值处理，得到与第二队列对应的一个第二特征信息。本实现方式中，通过对第二队列包括的所有支持样本的特征信息进行取平均值处理，得到与第二队列对应的一个第二特征信息，操作简单，易于实现。

在第一方面的一种可能实现方式中，在第一神经网络为用于回归的神经网络的情况下，方法还包括：训练设备将每个第一支持样本的标注结果添加至特征信息库的第一队列中。其中，特征信息库包括与 M 组支持样本一一对应的 M 个队列，第一队列为 M 个队列中与第一组对应的一个队列，用于存储与第一组的支持样本对应的特征信息。训练设备根据特

征信息库，获取与 M 组支持样本对应的第二特征信息，包括：训练设备对第二队列包括的所有支持样本的特征信息进行取平均值处理，得到与第二队列对应的一个第二特征信息。训练设备根据特征信息库，获取与 M 组支持样本对应的第一标注结果，包括：训练设备对第二队列包括的所有支持样本的标注结果进行取平均值处理，得到与第二队列对应的一个第一标注结果，第二队列为 M 个队列中任一队列。

在第一方面的一种可能实现方式中，训练设备可以预先设置第二队列的长度阈值为第一阈值，也即第二队列中最多只能存储第一阈值个支持样本的特征信息。

在第一方面的一种可能实现方式中，第二队列采用的为先进先出的队列结构。

在第一方面的一种可能实现方式中，训练设备根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一特征信息，生成与第一查询样本对应的第二预测结果，包括：训练设备根据第二特征信息和第一特征信息，执行相似度计算操作，得到第一相似度信息，第一相似度信息指示一个第一查询样本的第一特征信息与 M 个第二特征信息之间的相似度，第一查询样本与 M 组支持样本中每组支持样本之间的相似度。训练设备根据第一相似度信息和 M 个第一标注结果，生成第二预测结果。

本实现方式中，利用 M 组支持样本中每组支持样本的特征和与 M 组支持样本对应的 M 个第一标注结果，通过计算相似度的方式，生成第二预测结果，操作简单，易于实现。

在第一方面的一种可能实现方式中，在第一神经网络为对样本进行分类的神经网络的情况下，第一样本子集合包括 N 类第一支持样本，第一预测结果指示第一查询样本在 N 类中的预测类别。训练设备根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的特征信息，生成与第一查询样本对应的第二预测结果，包括：训练设备根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的特征信息，生成第一分类类别。其中，第一分类类别指示第一查询样本在 M 类中的预测类别，一个第一分类类别中包括的 M 个元素用于指示一个第一查询样本归属于 M 个类别中每个类别的概率。训练设备从第一分类类别中获取第二预测结果，第二预测结果指示第一查询样本在 N 类中的预测类别，第二预测结果可以表现为包括从前述 M 个元素中选取 N 个元素，选取的 N 个元素用于指示一个第一查询样本归属于 N 个类别中每个类别的概率。

本实现方式中，训练设备先根据 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的特征信息，生成第一查询样本归属于 M 个类别中每个类别的概率，再根据第一查询样本归属于 M 个类别中每个类别的概率确定第一查询样本归属于 N 个类别中每个类别的概率，从而保证在一个小样本任务中，训练设备在对一个第一查询样本进行分类的过程中，学习到了 M 个类别中每个类别的特征，以提高训练后的第一神经网络的输出结果的精度。

在第一方面的一种可能实现方式中，在第一神经网络为用于回归的神经网络的情况下，训练设备根据第一相似度信息和 M 个第一标注结果，生成第二预测结果，包括：训练设备根据第一相似度信息，生成 M 个第一标注结果中每个第一标注结果的权重，并根据 M 个第一标注结果中每个第一标注结果的权重和 M 个第一标注结果，进行加权求和，以生成第

二预测结果。

在第一方面的一种可能实现方式中，第一神经网络为对样本进行回归处理的神经网络。本实现方式中，训练后的第一神经网络不仅可以用于对样本进行分类，而且可以用于对样本进行回归，扩展了本方案的应用场景。

5 在第一方面的一种可能实现方式中，**M** 组的分组依据为以下中的一项或多项：与支持样本对应的用户、地点或时间。本实现方式中，提供了在回归这一应用场景下，**M** 组支持样本的多种分组依据，提高了本方案的实现灵活性。

在第一方面的一种可能实现方式中，第一样本子集合中包括的至少一个第一支持样本来自于 **M** 组的同一组，第一样本子集合中包括的至少一个第一查询样本来自于 **M** 组的同一组。
10

在第一方面的一种可能实现方式中，第一神经网络为以下中的任一种神经网络：全连接神经网络、卷积神经网络或残差神经网络。本实现方式中，第一神经网络可以为全连接神经网络、卷积神经网络或残差神经网络，提高了本方案的实现灵活性。

第二方面，本申请实施例还提供一种样本的处理方法，可用于人工智能领域的元学习领域中。方法包括：执行设备获取第三支持样本和第三查询样本；执行设备将第三支持样本和第三查询样本输入至第一神经网络，以通过第一神经网络输出的与第三查询样本对应的第三预测结果，第三预测结果为根据第三支持样本的特征信息和第三查询样本的特征信息之间的相似度生成的。其中，第一神经网络是根据第一损失函数和第二损失函数训练得到的，第一损失函数指示第一神经网络输出的预测结果和正确结果之间的相似度，第二损失函数指示第一神经网络输出的预测结果和与第三查询样本对应的第二预测结果之间的相似度，或者，第二损失函数指示第二预测结果和正确结果之间的相似度，第二预测结果为根据与 **M** 组支持样本对应的第二特征信息、与 **M** 组支持样本对应的正确结果和第三查询样本的特征信息得到的，第三支持样本能够归属于第三组的支持样本，第三查询样本能够归属于第三组的查询样本，第三组为 **M** 组中的一个组。
15
20

25 对于本申请实施例第二方面以及第二方面的各种可能实现方式中步骤的具体实现方式、每种可能实现方式中名词的具体含义，以及每种可能实现方式所带来的有益效果，均可以参考第一方面中各种可能的实现方式中的描述，此处不再一一赘述。

第三方面，本申请实施例提供一种神经网络的训练装置，可用于人工智能领域的元学习领域中。神经网络的训练装置包括获取模块、输入模块、生成模块和训练模块。其中，
30 获取模块，用于从样本集合中获取第一样本子集合，样本集合包括 **M** 组支持样本和 **M** 组查询样本，第一样本子集合包括第一支持样本和第一查询样本；输入模块，用于将第一样本子集合输入至第一神经网络，第一神经网络生成第一查询样本的第一特征信息，并得到第一神经网络输出的与第一查询样本对应的第一预测结果，第一预测结果为根据第一支持样本的特征信息和第一特征信息之间的相似度生成的；生成模块，用于根据与 **M** 组支持样本对应的第二特征信息、与 **M** 组支持样本对应的第一标注结果和第一查询样本的第一特征信息，生成与第一查询样本对应的第二预测结果，第一标注结果中的每一个均指示与 **M** 组支持样本中一组支持样本对应的正确结果；训练模块，用于根据第一损失函数和第二损失
35

函数，对第一神经网络进行迭代训练，直至满足预设条件。第一损失函数指示第一预测结果和第二标注结果之间的相似度，第二损失函数指示第一预测结果与第二预测结果之间的相似度，或者，第二损失函数指示第二预测结果和第二标注结果之间的相似度，第二标注结果为第一查询样本的正确结果。

5 本申请实施例第三方面中，神经网络的训练装置还可以用于实现第一方面各种可能实现方式中训练设备执行的步骤，对于本申请实施例第三方面以及第三方面的各种可能实现方式中某些步骤的具体实现方式，以及每种可能实现方式所带来的有益效果，均可以参考第一方面中各种可能的实现方式中的描述，此处不再一一赘述。

10 第四方面，本申请实施例提供一种样本处理装置，可用于人工智能领域的元学习领域中。样本处理装置包括获取模块和输入模块。其中，获取模块，用于获取第三支持样本和第三查询样本；输入模块，用于将第三支持样本和第三查询样本输入至第一神经网络，以通过第一神经网络输出的与第三查询样本对应的第三预测结果，第三预测结果为根据第三支持样本的特征信息和第三查询样本的特征信息之间的相似度生成的。第一神经网络是根据第一损失函数和第二损失函数训练得到的，第一损失函数指示第一神经网络输出的预测结果和正确结果之间的相似度，第二损失函数指示第一神经网络输出的预测结果和与第三查询样本对应的第二预测结果之间的相似度，或者，第二损失函数指示第二预测结果和正确结果之间的相似度，第二预测结果为根据与M组支持样本对应的第二特征信息、与M组支持样本对应的正确结果和第三查询样本的特征信息得到的，第三支持样本能够归属于第三组的支持样本，第三查询样本能够归属于第三组的查询样本，第三组为M组中的一个组。

20 本申请实施例第四方面中，神经网络的训练装置还可以用于实现第二方面各种可能实现方式中执行设备执行的步骤，对于本申请实施例第四方面以及第四方面的各种可能实现方式中某些步骤的具体实现方式，以及每种可能实现方式所带来的有益效果，均可以参考第二方面中各种可能的实现方式中的描述，此处不再一一赘述。

25 第五方面，本申请实施例提供了一种训练设备，可以包括处理器，处理器和存储器耦合，存储器存储有程序指令，当存储器存储的程序指令被处理器执行时实现上述第一方面所述的神经网络的训练方法。对于处理器执行第一方面的各个可能实现方式中训练设备执行的步骤，具体均可以参阅第一方面，此处不再赘述。

30 第六方面，本申请实施例提供了一种执行设备，可以包括处理器，处理器和存储器耦合，存储器存储有程序指令，当存储器存储的程序指令被处理器执行时实现上述第二方面所述的样本的处理方法。对于处理器执行第二方面的各个可能实现方式中执行设备执行的步骤，具体均可以参阅第二方面，此处不再赘述。

第七方面，本申请实施例提供了一种计算机可读存储介质，所述计算机可读存储介质中存储有计算机程序，当其在计算机上运行时，使得计算机执行上述第一方面所述的神经网络的训练方法，或者，使得计算机执行上述第二方面所述的样本的处理方法。

35 第八方面，本申请实施例提供了一种电路系统，所述电路系统包括处理电路，所述处理电路配置为执行上述第一方面所述的神经网络的训练方法，或者，所述处理电路配置为执行上述第二方面所述的样本的处理方法。

第九方面，本申请实施例提供了一种计算机程序，当其在计算机上运行时，使得计算机执行上述第一方面所述的神经网络的训练方法，或者，使得计算机执行上述第二方面所述的样本的处理方法。

第十方面，本申请实施例提供了一种芯片系统，该芯片系统包括处理器，用于支持训练设备或执行设备实现上述方面中所涉及的功能，例如，发送或处理上述方法中所涉及的数据和/或信息。在一种可能的设计中，所述芯片系统还包括存储器，所述存储器，用于保存服务器或通信设备必要的程序指令和数据。该芯片系统，可以由芯片构成，也可以包括芯片和其他分立器件。

附图说明

- 图 1 为本申请实施例提供的人工智能主体框架的一种结构示意图；
- 图 2 为本申请实施例提供的样本处理系统的一种系统架构图；
- 图 3 为本申请实施例提供的神经网络的训练方法的一种流程示意图；
- 图 4 为本申请实施例提供的神经网络的训练方法中第一样本子集合的一种示意图；
- 图 5 为本申请实施例提供的神经网络的训练方法中第一样本子集合的另一种示意图；
- 图 6 为本申请实施例提供的神经网络的训练方法中生成第一预测结果的一种流程示意图；
- 图 7 为本申请实施例提供的神经网络的训练方法中特征信息库的一种示意图；
- 图 8 为本申请实施例提供的神经网络的训练方法的另一种流程示意图；
- 图 9 为本申请实施例提供的神经网络的训练方法的又一种流程示意图；
- 图 10 为本申请实施例提供的样本的处理方法的一种流程示意图；
- 图 11 为本申请实施例提供的神经网络的训练装置的一种结构示意图；
- 图 12 为本申请实施例提供的神经网络的训练装置的另一种结构示意图；
- 图 13 为本申请实施例提供的样本处理装置的一种结构示意图；
- 图 14 为本申请实施例提供的训练设备的一种结构示意图；
- 图 15 为本申请实施例提供的执行设备的一种结构示意图；
- 图 16 为本申请实施例提供的芯片的一种结构示意图。

具体实施方式

本申请实施例提供了一种用于神经网络的训练方法以及相关设备，在一个小样本任务的训练过程中不仅学习到一个小样本任务中支持样本的特征，而且学习到 **M** 组支持样本的特征，也即神经网络在一个小样本学习中不仅能够学习到一个小样本任务内的相对关系，而且能够学习到多个小样本任务之间的绝对关系，从而提高输出的预测结果的精度。

本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的术语在适当情况下可以互换，这仅仅是描述本申请的实施例中相同属性的对象在描述时所采用的区分方式。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，以便

包含一系列单元的过程、方法、系统、产品或设备不必限于那些单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它单元。

下面结合附图，对本申请的实施例进行描述。本领域普通技术人员可知，随着技术的发展和场景的出现，本申请实施例提供的技术方案对于类似的技术问题，同样适用。

5 首先对人工智能系统总体工作流程进行描述，请参见图 1，图 1 示出的为人工智能主体框架的一种结构示意图，下面从“智能信息链”（水平轴）和“IT 价值链”（垂直轴）两个维度对上述人工智能主题框架进行阐述。其中，“智能信息链”反映从数据的获取到处理的一系列过程。举例来说，可以是智能信息感知、智能信息表示与形成、智能推理、智能决策、智能执行与输出的一般过程。在这个过程中，数据经历了“数据—信息—知识—智慧”
10 的凝练过程。“IT 价值链”从人工智能的底层基础设施、信息（提供和处理技术实现）到系统的产业生态过程，反映人工智能为信息技术产业带来的价值。

（1）基础设施

基础设施为人工智能系统提供计算能力支持，实现与外部世界的沟通，并通过基础平台实现支撑。通过传感器与外部沟通；计算能力由智能芯片提供，作为示例，该智能芯片
15 包括中央处理器（central processing unit, CPU）、神经网络处理器（neural-network processing unit, NPU）、图形处理器（graphics processing unit, GPU）、专用集成电路（application specific integrated circuit, ASIC）、现场可编程逻辑门阵列（field programmable gate array, FPGA）等硬件加速芯片；基础平台包括分布式计算框架及网络等相关的平台保障和支持，可以包括云存储和计算、互联互通网络等。举例来说，传感器和外部沟通获取数据，这些数据提供
20 供给基础平台提供的分布式计算系统中的智能芯片进行计算。

（2）数据

基础设施的上一层的数据指示人工智能领域的数据来源。数据涉及到图形、图像、语音、文本，还涉及到传统设备的物联网数据，包括已有系统的业务数据以及力、位移、液位、温度、湿度等感知数据。

25 （3）数据处理

数据处理通常包括数据训练，机器学习，深度学习，搜索，推理，决策等方式。

其中，机器学习和深度学习可以对数据进行符号化和形式化的智能信息建模、抽取、预处理、训练等。

推理是指在计算机或智能系统中，模拟人类的智能推理方式，依据推理控制策略，利用形式化的信息进行机器思维和求解问题的过程，典型的功能是搜索与匹配。
30

决策是指智能信息经过推理后进行决策的过程，通常提供分类、排序、预测等功能。

（4）通用能力

对数据经过上面提到的数据处理后，进一步基于数据处理的结果可以形成一些通用的能力，比如可以是算法或者一个通用系统，例如，图像的分类、图像的个性化管理、电池
35 充电个性化管理、文本分析、计算机视觉的处理、语音识别等等。

（5）智能产品及行业应用

智能产品及行业应用指人工智能系统在各领域的产品和应用，是对人工智能整体解决

方案的封装，将智能信息决策产品化、实现落地应用，其应用领域主要包括：智能终端、智能制造、智能交通、智能家居、智能医疗、智能安防、自动驾驶、平安城市等。

本申请实施例主要应用于采用元学习技术对神经网络进行训练的应用场景中，训练后的神经网络可以应用上述各种应用领域中以实现分类、回归或其他功能，训练后的神经网络的处理对象可以为图像样本、离散数据样本、文本样本或语音样本等。作为示例，例如在图像分类场景中，由于待分类对象的种类繁多，训练设备中一般不能保证每种类别的对象都存在大量的训练样本，则可以采用元学习技术来训练神经网络学会学习（learning to learn）的能力，也即训练神经网络学会学习新类别图像的能力。作为另一示例，例如在图像的个性化管理场景中，智能终端在为用户拍照时，可以连续拍多张图像，从多张图像中选取最符合用户审美的图像展示给用户，由于不同用户的审美不同，而训练设备中一般也无法采集到每个用户的大量的训练样本，则可以元学习技术来训练神经网络，以训练神经网络学会学习新用户对图像的审美。作为另一示例，例如在电池充电个性化管理场景中，可以根据用户的手机型号、电池型号、充电地点、充电时刻或其他数据的离散数据，回归出充电时长，不同用户的离散数据与充电时长之间的回归逻辑可以不同，训练设备中一般无法采集到每个用户的大量的训练样本，则可以采用元学习技术来训练神经网络，以训练神经网络学会学习新用户的离散数据和充电时长之间的逻辑关系。此外，本申请实施例还可以应用于语音唤醒个性化管理场景、屏下指纹个性化管理场景等。应当理解，此处举例仅为方便对本申请实施例的应用场景进行理解，不对本申请实施例的应用场景进行穷举。

基于上述种种场景中，本申请实施例提供了一种神经网络的训练方法。为了便于理解本方案，首先结合图 2 对本申请实施例提供的神经网络的训练方法所在的样本处理系统的系统架构进行介绍，请先参阅图 2，图 2 为本申请实施例提供的样本处理系统的一种系统架构图。在图 2 中，样本处理系统 200 包括执行设备 210、训练设备 220、数据库 230 和数据存储系统 240，执行设备 210 中包括计算模块 211。

其中，在训练阶段，数据库 230 用于存储有用于训练的样本集合，样本集合可以划分为支持样本集（support set）和查询样本集（query set），支持样本集中包括 M 组支持样本和每个支持样本的标注结果，查询样本集中包括 M 组查询样本和每个查询样本的标注结果，M 的取值为大于或等于 1 的整数。训练设备 220 生成用于处理样本的目标模型/规则 201，并利用数据库中的样本集合对目标模型/规则 201 进行迭代训练，得到成熟的目标模型/规则 201，该目标模型/规则 201 具体表现为神经网络。训练设备 220 得到的神经网络可以应用不同的系统或设备中。

在推理阶段，执行设备 210 可以调用数据存储系统 240 中的数据、代码等，也可以将数据、指令等存入数据存储系统 240 中。数据存储系统 240 可以置于执行设备 210 中，也可以为数据存储系统 240 相对执行设备 210 是外部存储器。计算模块 211 可以通过神经网络对执行设备 210 获取到的样本进行处理，得到预测结果，预测结果的具体表现形式与神经网络的功能相关。

本申请的一些实施例中，例如图 2 中，“用户”可以直接与执行设备 210 进行交互，也即执行设备 210 与客户设备集成于同一设备中。但图 2 仅是本发明实施例提供的两种样本

处理系统的架构示意图，图中所示设备、器件、模块等之间的位置关系不构成任何限制。在本申请的另一些实施例中，执行设备 210 和客户设备可以为分别独立的设备，执行设备 210 配置有输入/输出接口，与客户设备进行数据交互，“用户”可以通过客户设备向输入/输出接口输入采集到的样本，执行设备 210 通过输入/输出接口将预测结果返回给客户设备。

5 结合上述描述，本申请实施例提供了一种神经网络的训练方法，在神经网络的训练阶段，在执行每个小样本任务的过程中，不仅通过的传统方式对神经网络进行训练，还会根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的第一特征信息，生成与第一查询样本对应的第二预测结果，第一标注结果中的每一个均指示与 M 组支持样本中一组支持样本对应的正确结果。并增设第二损失函数，第二损失函数的训练目标为拉近第一预测结果与第二预测结果之间的相似度，或者，第二损失函数的训练目标为拉近第二预测结果和第二标注结果之间的相似度。从而在与一个小样本任务对应的训练过程中，不仅能够学习到一个小样本任务中支持样本的特征，而且能够学习到 M 组支持样本的特征，也即神经网络在一个小样本学习中不仅能够学习到一个小样本任务内的相对关系，而且能够学习到多个小样本任务之间的绝对关系，从而提高输出的预测结果的精度。由于该方法对训练阶段和推理阶段均有影响，而训练阶段和推理阶段的实现流程有所不同，下面分别对前述两个阶段的具体实现流程进行描述。

一、训练阶段

本申请实施例中，训练阶段指的是上述图 2 中训练设备 220 对目标模型/规则 201 执行训练操作的过程。由于通过本申请实施例训练得到的神经网络可以用于分类、回归或其他功能，而当神经网络用于分类和当神经网络用于回归时，训练的步骤有所不同，以下分别对两种功能的训练过程进行介绍。

(1) 训练的用于分类的神经网络

本申请实施例中，训练的用于分类的神经网络指的是神经网络输出的预测结果是基于分类网络输出的，作为示例，例如语音识别、语音翻译等均为通过分类网络实现的，也视为训练的用于分类的神经网络。具体的，请参阅图 3，图 3 为本申请实施例提供的神经网络的训练方法的一种流程示意图，方法可以包括：

301、训练设备创建特征信息库。

本申请实施例中，训练设备在对神经网络进行训练之前，会创建特征信息库。特征信息库又称为元存储库 (meta-repository)，特征信息库包括与 M 组支持样本一一对应的 M 个队列。本实施例中 M 组支持样本中不同组的支持样本的类别不同，也即 M 组支持样本共包括 M 类支持样本。 M 个队列中的每个队列均用于存储与一组支持样本对应的特征信息和与一组支持样本对应的标注结果。

由于本实施例中神经网络的功能为进行分类，则 M 个队列中不同队列的分类依据为类别。作为示例，例如神经网络是用于进行图像分类，则不同类别可以为猫、狗、海豚、河马等。作为再一示例，例如神经网络是用于进行图像个性化管理中，则不同类别可以为用户甲喜欢的图像、用户甲不喜欢的图像、用户乙喜欢的图像、用户乙不喜欢的图像等。或者，不同类别可以为甲类用户喜欢的图像、甲类用户不喜欢的图像、乙类用户喜欢的图像、

乙类用户不喜欢的图像等。作为再一示例，例如神经网络是用于进行语音的个性化识别，则不同类别的语音之间的区别可以为口音不同，应理解，此处不做穷举。

其中， M 个队列中每个队列均可以采用键-值 (key-value) 的方式存储。训练设备通过不同的“键”来区别不同的队列，键也可以称为令牌，用于唯一标识一个队列。 M 个队列中每个队列还可以包括队列的类别。可选地，由于本实施例中不同队列的类别不同，则可以直接将每个队的类别作为该队的标识信息（也即该队的键）。进一步地，每个队的类别可以表现为独热 (one-hot) 编码的形式。每个队列中的“值”包括多个支持样本的特征信息。用于指示类别的独热编码具体可以表现为一个包括 M 个元素的向量。作为示例，例如 M 的取值为 64，64 种类别分别为猫、狗、狮子、草莓、沙发、...、茶几和苹果，独热编码为 (0, 0, 1, 0, 0, ..., 0, 0) 代表的类别为狮子，此处未对 64 种类别进行穷举，独热编码中未展开示出的元素均为 0，应理解，在其他实现方式中， M 的取值也可以为更多或更少，此处举例仅为方便理解，不用于限定本方案。

具体的，由于样本集合中共包括 M 个队列，则训练设备需要初始化 M 个键（也可以称为初始化 M 个令牌）。可选地，训练设备在初始化阶段可以为每个队列设置队列长度阈值，也即每个队列中存储的特征信息的个数需要小于或等于第一阈值。第一阈值的取值可以为一个小样本任务中每类支持样本的数量与每列查询样本的数量之和的预设倍数。作为示例，例如预设倍数可以为 5、6、7、8、9 或 10 等，第一阈值的取值可以为 500、1000 或其他数值等，此处不做限定。为更直观的理解本方案， M 个队列中一个队列的表示可以如下：

$$Q_c = \{\phi(x_i), y_i = c\} \text{ and } |Q_c| = \max_len; (1)$$

其中， Q_c 代表 M 个队列中类别为 c 的队列， $\phi(x_i), y_i = c$ 代表类别为 c 的队列用于存储标注类别为 c 的支持样本的特征信息， $|Q_c| = \max_len$ 代表类别为 c 的队列的长度阈值为 $\max_len$ ，也即一个队列中最多可以存储 $\max_len$ 个支持样本的特征信息， $\max_len$ （也即第一阈值）的数值为预先设定好的，类别 c 为 M 个类别中的任一个类别，应理解，式 (1) 中的举例仅为方便理解本方案，不用于限定本方案。

进一步可选地， M 个队列中每个队列可以采用先近先出 (first in first out, FIFO) 的队列结构。

302、训练设备从样本集合中获取第一样本子集合，第一样本子集合包括第一支持样本和第一查询样本。

本申请实施例中，训练设备在执行小样本任务之前，从样本集合中获取第一样本子集合，一个第一样本子集合对应一个小样本任务。样本集合中共包括 M 组支持样本和 M 组查询样本， M 组支持样本和 M 组查询样本中包括的样本互相不重合，以保证在一个小样本任务中支持样本和查询样本的不重合。

其中，一个第一样本子集合中包括 N 类第一支持样本和 n 个第一查询样本， N 类第一支持样本中的每类第一支持样本中均包括至少一个第一支持样本，任一个查询样本的正确

类别均包含于前述 N 类类别中； N 为大于或等于 1 且小于或等于 M 的整数， n 为大于或等于 1 的整数。其中， M 组支持样本和 M 组查询样本的类别均可参阅步骤 301 中对 M 个队列的类别的介绍，此处不做赘述。

具体的，在一种应用场景下，神经网络用于对样本中的对象进行分类， N 的取值为大于或等于 2。则步骤 302 可以包括：训练设备从 M 组支持样本中随机选取 N 类第一支持样本，并从 M 组查询样本包括的 N 类查询样本中随机选取 n 个第一查询样本， n 个第一查询样本中的任一个第一查询样本均归属于前述 N 类中的一类。作为示例，例如 N 的取值为 5，支持样本和查询样本具体为图像， N 类支持图像的类别分别为猫、狗、狮子、草莓和苹果，则任一个查询图像的类别为猫、狗、狮子、草莓或苹果中的一种。作为另一示例，例如 N 的取值为 3，支持样本和查询样本具体为语音，则 N 类支持语音的类别分别为机主本人的语音、甲用户（非机主）的语音和乙用户（非机主）的语音，应理解，此处举例仅为方便理解本方案，不用于限定本方案。应理解，此处举例仅为方便理解本方案。

为更直观的理解第一样本子集合的概念，请参阅图 4，图 4 为本申请实施例提供的神经网络的训练方法中第一样本子集合的一种示意图。图 4 中以支持样本和查询样本均为图像为例。如图 4 所示，用于执行一个小样本任务的第一样本子集合中包括 4 类第一支持图像和多个第一查询图像；4 个类别分别为狗、狼、桌子和沙发，每类第一支持图像中包括至少一个第一支持图像，多个第一查询图像中包括正确类别为狗和沙发的图像，类别狗和类别沙发均在前述 4 个类别中，应理解，图 4 中的举例仅为方便理解第一样本子集合的概念，不用于限定本方案。

在另一种应用场景下，神经网络用于对图像进行个性化管理， N 的取值为 2。则样本集合中 M 类（也即对应 M 组）支持样本和 M 类查询样本中不同类别的样本之间分类依据和用户有关。在一种情况下，前述 M 类的分类依据为用户和是否喜欢两个维度，也即一组样本中包括一个用户喜欢的图像，或者，包括一个用户不喜欢的图像。可选地，训练设备可以根据样本集合中多个用户中每个用户的数据，得到每个用户的特征，然后基于每个用户的特征对多个用户进行聚类，则一组样本中包括一类用户喜欢的图像，或者，包括一类用户不喜欢的图像。

步骤 302 可以包括：训练设备从与 M 类支持图像包括的多个用户随机确定一个第一用户，或者，从 M 类支持图像包括的多类用户中随机确定一类第一类用户，也即从 M 组支持图像中确定了两组支持图像，分别为第 P 组支持图像和第 Q 组支持图像，从 M 组查询图像中确定了两组查询图像，分别为第 P 组查询图像和第 Q 组查询图像。其中，第 P 组支持图像中包括第一用户（或第一类用户）喜欢的支持图像，第 Q 组支持图像包括第一用户（或第一类用户）不喜欢的支持图像；第 P 组查询图像中包括第一用户（或第一类用户）喜欢的查询图像，第 Q 组查询图像包括第一用户（或第一类用户）不喜欢的查询图像。

训练设备从第 P 组支持图像中随机选取至少一个支持图像，并从第 Q 组支持图像中随机选取至少一个支持图像，将前述选取的支持图像作为第一样本子集合中的第一支持图像。训练设备从第 P 组查询图像中随机选取至少一个查询图像，和/或从第 Q 组查询图像中随机选取至少一个查询图像，将前述选取的查询图像作为第一样本子集合中的第一查询图像。

为更直观的理解第一样本子集合的概念，请参阅图 5，图 5 为本申请实施例提供的神经网络的训练方法中第一样本子集合的一种示意图。图 5 中以支持样本和查询样本均为图像为例，第一样本集合中包括的所有图像均来选自于用户甲的图像。如图 5 所示，第一样本子集合中包括两类第一支持图像，分别为用户甲不喜欢的图像和用户甲喜欢的图像，第一样本子集合中包括多个第一查询图像，如图 5 所示，包括用户甲不喜欢的图像和用户甲喜欢的图像，应理解，在实际情况中，第一样本子集合包括的第一查询图像中也可以只选取用户甲喜欢的图像，或者只选取用户甲不喜欢的图像，图 5 中的举例仅为方便理解第一样本子集合的概念，不用于限定本方案。

进一步地，训练设备还会获取每个第一支持样本的正确类别（也即每个第一支持样本的标注结果）以及每个第一查询样本的正确类别（也即每个第一查询样本的标注结果）。其中，在一个小样本任务中，支持样本用于引导神经网络学习 N 类第一支持样本中每类第一支持样本的特征，查询样本用于测试神经网络的分类能力。

303、训练设备将第一样本子集合输入至第一神经网络，获取第一神经网络生成第一查询样本的第一特征信息和第一支持样本的特征信息，并获取第一神经网络输出的与第一查询样本对应的第一预测结果，第一神经网络用于对样本进行分类。

本申请实施例中，训练设备在获取到用于执行一个小样本任务的第一样本子集合之后，可以将 N 类第一支持样本、每个第一支持样本的标注结果（也即正确类别）和 n 个第一查询样本输入至第一神经网络中，以通过第一神经网络对每个第一支持样本和每个第一查询样本执行特征提取操作，得到第一查询样本的第一特征信息和每个第一支持样本的特征信息。

训练设备通过第一神经网络根据 N 类第一支持样本中每个第一支持样本的特征信息，分别生成 N 类中每个类别的类中心。具体的，针对 N 类中的第二类，第二类为 N 类中的任意一类。训练设备在通过第一神经网络生成每个第一支持样本的特征信息之后，可以根据每个第一支持样本的标注结果，可以获取到归属于第二类的每个第一支持样本的特征信息，训练设备通过第一神经网络对归属于第二类的每个第一支持样本的特征信息进行求平均值，得到 N 类中第二类的类中心；训练设备通过第一神经网络对 N 类中每一类均执行前述操作，得到 N 类中每一类的类中心。

训练设备在得到 N 类中每一类的类中心之后，通过第一神经网络生成与每个第一查询样本对应的第一预测结果，第一预测结果为根据第一支持样本的特征信息和第一查询样本的第一特征信息之间的相似度生成的。

针对至少一个第一查询样本中的任一个第一查询样本，训练设备通过第一神经网络计算一个第一查询样本的第一特征信息与 N 类中每个类别的类中心之间的距离，也即计算一个第一查询样本的第一特征信息与 N 类中每个类别的类中心之间的相似度，得到该一个第一查询样本的第二分类类别（也即该一个第一查询样本的第一预测结果）。其中，第一查询样本的第二分类类别也可以称为第一查询样本的标签，第一查询样本的第一特征信息与 N 类中某一类别的类中心距离越近，则第一查询样本为该类别的概率越大。训练设备通过第一神经网络对每个第一查询样本均执行前述操作，以得到每个第一查询样本的第二分类类

别（也即每个第一查询样本的第一预测结果），第一查询样本的第二分类类别用于指示第一查询样本在 N 类中的预测分类。

作为示例，例如第一查询样本为图像， N 类分别为狗、狼、桌子和沙发，第一查询图像的第二分类类别指示第一查询图像的分类是狗。作为另一示例，例如第一查询样本为图像， N 类分别为用户甲喜欢的图像和用户甲不喜欢的图像，第一查询图像的第二分类类别指示第一查询图像的分类是户甲喜欢的图像。

其中，第一神经网络具体可以为卷积神经网络或残差神经网络，作为示例，例如第一神经网络可以为 4 层的卷积神经网络，作为另一示例，例如第一神经网络可以为残差网络（residual network, ResNet）12 等，第一神经网络还可以为其他用于进行分类的神经网络等，此处不做穷举。第一神经网络包括特征提取网络和第一分类网络，第一分类网络可以具体表现为分类器。特征提取网络用于生成第一支持样本的特征信息和第一查询样本的特征信息，第一分类网络用于生成第一查询样本的类别指示信息。

为更直观的理解本方案，请参阅图 6，图 6 为本申请实施例提供的神经网络的训练方法中生成第一预测结果的一种流程示意图。图 6 中以支持样本和查询样本均为图像、 N 的取值为 4，4 个类别分别为狗、狼、桌子和沙发为例。训练设备将第一样本子集输入第一神经网络中，通过第一神经网络对每个第一支持图像进行特征提取，并根据每个第一支持图像的特征信息生成 4 个类中心，图 6 中以 4 个三角形代表 4 个类中心，分别指示类别为狗的所有支持图像的特征、类别为狼的所有支持图像的特征、类别为桌子的所有支持图像的特征和类别为沙发的所有支持图像的特征。训练设备通过第一神经网络生成每个第一查询图像的特征信息，并通过第一神经网络计算一个第一查询图像的特征信息与 4 个类中心之间的距离，以输出该一个第一查询样本的第一预测结果，前述第一预测结果指示该一个第一查询样本在 4 个类别中的预测类别，应理解，图 4 中的举例仅为方便理解第一样本子集的概念，不用于限定本方案。

304、训练设备将第一支持样本的特征信息添加至特征信息库的第一队列中。

本申请实施例中，训练设备在每次通过步骤 303 生成了每个第一支持样本的特征信息之后，都会进行元存储库更新（meta repo update）。针对一个第一支持样本，在一种实现方式中，训练设备根据该一个第一支持样本对应的组别，从 M 个队列中选取与该一个第一支持样本对应的第一队列，将该一个第一支持样本的特征信息放入第一队列中。具体的，由于 M 个队列与 M 组支持样本一一对应，该一个第一支持样本选自于（choosed from）第一组，则将该第一支持样本的特征信息放入与第一组对应的第一队列，第一队列用于存储与第一组的支持样本对应的特征信息。

在另一种实现方式中，由于 M 个队列中每个队列都包括类别信息，训练设备根据该一个第一支持样本对应的标注结果（也即该一个第一支持样本的正确类别）从 M 个队列中选取与该一个第一支持样本对应的第一队列，将该一个第一支持样本的特征信息放入第一队列中。

可选地，由于第一队列采用的是“先进先出”的队列结构，则训练设备可以将该一个第一支持样本的特征信息添加至第一队列的队尾。

进一步可选地，由于可以预先将第一队列的长度设置为第一阈值，则训练设备在将第一支持样本的特征信息添加至第一队列之前，可以判断第一队列的长度是否达到第一阈值，过达到第一阈值，则将第一队列中位于队列头的特征信息删除，并将将该一个第一支持样本的特征信息添加至第一队列的队尾。

5 训练设备对第一样本子集合中每个第一支持样本执行前述操作，以将每个第一支持样本的特征信息添加至特征信息库的第一队列中。需要说明的是，由于一个第一样本子集合中包括 N 类第一支持样本，则训练设备会将第一样本子集合中所有第一支持样本的特征信息分别添加至 N 个不同的第一队列中。

10 为更直观的理解本方案，请参阅图 7，图 7 为本申请实施例提供的神经网络的训练方法中特征信息库的一种示意图。特征信息库中包括 M 个队列，图 7 中示出的 M 行分别对应 M 个队列，第二队列是 M 个队列中的任一个队列。如图 7 所示，第二队列中包括令牌和多个 $A1$ ，令牌也可以称为键，第二队列中的多个 $A1$ 为第二队列的值，在第一神经网络用于分类这一应用场景中，任一个 $A1$ 均指示一个支持样本的特征信息，图 7 中以一个队列中最多存储 12 个支持样本的特征信息为例。应理解，在实际情况下，在特征信息库初始时，每个队列中的值均为空的，是在对第一神经网络不断迭代更新的过程中不断补充每个队列的值，图 7 中的举例仅为方便理解特征信息库的概念，当应用于回归场景中时，每个 $A1$ 不仅指示一个支持样本的特征信息，还指示一个支持样本的标注结果，此处举例不用于限定本方案。

20 305、训练设备根据特征信息库，获取与 M 组支持样本对应的第二特征信息和与 M 组支持样本对应的第一标注结果。

本申请实施例中，训练设备根据特征信息库，获取与 M 组支持样本对应的 M 个第二特征信息和与 M 组支持样本对应的 M 个第一标注结果， M 个第一标注结果中的每一个均指示与 M 组支持样本中一组支持样本对应的正确结果（也即正确类别）。

25 特征信息库中包括与 M 组支持样本一一对应的 M 个队列，每个队列中可以包括一个或多个支持样本的特征信息。 M 个第二特征信息中每个第二特征信息均用于反映一个队列中所有支持样本的特征信息，由于 M 个队列中的一个队列对应于一个类别的支持样本，则与 M 组支持样本中一组支持样本对应的一个第二特征信息也可以理解为与 M 组支持样本中一组支持样本对应的一个类中心。

30 具体的，针对与 M 组支持样本中任一组支持样本对应的一个第二特征信息。训练设备从特征信息库包括的 M 个队列中获取第二队列包括的所有支持样本的特征信息和第二队列的类别，对第二队列包括的所有支持样本的特征信息进行取平均值处理，得到与第二队列对应的一个第二特征信息（也即得到与 M 组支持样本中一组支持样本对应的一个第二特征信息），并将第二队列的类别确定为与 M 组支持样本中一组支持样本对应的一个第一标注结果。其中，第二队列为 M 个队列中的任一个队列。

35 本申请实施例中，通过对第二队列包括的所有支持样本的特征信息进行取平均值处理，得到与第二队列对应的一个第二特征信息，操作简单，易于实现。

需要说明的是，本申请实施例不限定步骤 305 的执行顺序，步骤 305 可以在步骤 302、

步骤 303 和步骤 304 中任一个步骤之前或之后执行。

本申请实施例中，训练设备中配置有特征信息库，特征信息库中包括与 M 组支持样本一一对应的 M 个队列， M 个队列中每一个均用于存储一组支持样本的特征信息；在每次执行每次训练的过程中，会通过第一神经网络生成每个第一支持样本的特征信息，并将第一神经网络生成的每个第一支持样本的特征信息添加至对应的队列中。训练设备能够根据特征信息库，获取与 M 组支持样本对应的 M 个第二特征信息，也即训练设备在执行每次训练的过程中，会不断更新特征信息库。通过前述方式，训练设备利用每次训练过程中生成的特征信息，来生成 M 组中每组支持样本的特征，提高了训练过程的效率。

306、训练设备根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的第一特征信息，生成与第一查询样本对应的第二预测结果。

本申请实施例中，针对第一样本子集合包括的一个或多个第一查询样本中的任一个第一查询样本。训练设备在获取到与 M 组支持样本对应的 M 个第二特征信息、与 M 组支持样本对应的 M 个第一标注结果之后，根据 M 个第二特征信息和一个第一查询样本的第一特征信息，执行相似度计算操作，得到第一相似度信息，第一相似度信息指示一个第一查询样本的第一特征信息与 M 个第二特征信息之间的相似度，也即指示第一查询样本与 M 组支持样本中每组支持样本之间的相似度。训练设备根据第一相似度信息和与 M 组支持样本对应的 M 个第一标注结果，生成第二预测结果。

本申请实施例中，利用第一查询样本的特征与 M 组支持样本中每组支持样本的特征之间的相似度和与 M 组支持样本对应的 M 个第一标注结果，生成第一查询样本的第二预测结果，通过计算相似度的方式，生成第二预测结果，操作简单，易于实现。

具体的，训练设备根据第一相似度信息和与 M 组支持样本对应的 M 个第一标注结果，成第一分类类别。其中，第一分类类别指示该一个第一查询样本在 M 类中的预测类别； M 个第二特征信息与 M 个类别一一对应，第三特征信息为 M 个第二特征信息中的任一个特征信息，第一特征信息与第三特征信息的相似度越高，第一查询样本归属于与第三特征信息对应的类别的概率越大。

训练设备从第一分类类别中获取第二预测结果，第二预测结果指示第一查询样本在 N 类中的预测类别。其中，一个第一分类类别可以表现为包括 M 个元素的向量，一个第一分类类别中包括的 M 个元素用于指示一个第一查询样本归属于 M 个类别中每个类别的概率。第二预测结果可以表现为包括从前述 M 个元素中选取 N 个元素，选取的 N 个元素用于指示一个第一查询样本归属于 N 个类别中每个类别的概率。

作为示例，例如 M 的取值为 6， M 个类别分别为用户甲喜欢的图像、用户甲不喜欢的图像、用户乙喜欢的图像、用户乙不喜欢的图像、用户丙喜欢的图像和用户丙不喜欢的图像， N 的取值为 2， N 个类别分别为用户乙喜欢的图像和用户乙不喜欢的图像。第一分类类别具体表现为 (0.05, 0.1, 0.7, 0.01, 0.09, 0.05)，指示一个第一查询图像归属于 M 个类别中每个类别的概率；第二预测结果可以为 (0.7, 0.01)，指示一个第一查询图像归属于 N 个类别中每个类别的概率。

作为另一示例，例如 M 的取值为 12， N 的取值为 3， M 个类别分别为猫、狗、狮子、

草莓、沙发、茶几、苹果、栗子、牛奶、狼、老虎和海豚，N 个类别分别为猫、狗和狮子，第一分类类别具体表现为 (0.01, 0.8, 0.01, 0.01, 0.02, 0.05, 0.01, 0.01, 0.01, 0.02, 0.03, 0.02)，指示一个第一查询图像归属于 M 个类别中每个类别的概率；第二预测结果可以为 (0.01, 0.8, 0.01)，指示一个第一查询图像归属于 N 个类别中每个类别的概率。

5 应理解，M 还可以取更大或更小的数值，N 还可以取更大的数值，上述种种举例仅为方便理解本方案，不用于限定本方案。

具体的，训练设备将与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和至少一个第一查询样本中每个第一查询样本的第一特征信息输入至第二分类网络中，以通过第二分类网络生成第一分类类别，再从第一分类类别中获取第二预测结果。其中，第二分类网络与第一分类网络的区别在于，第二分类网络能够实现 M 个类别的分类，第一分类网络仅能实现 N 个类别的分类。第二分类网络可以表现为分类器，进一步地，第二分类网络可以表现为线性分类器，作为示例，例如第二分类网络具体表现为支持向量机 (support vector machine, SVM)。第一神经网络中的特征提取网络和第二分类网络可以组成第二神经网络，进一步地，第二神经网络的公式可以如下：

15
$$\hat{h} = \text{SVM} \circ \phi; (2)$$

其中， $\hat{h}$ 代表第二神经网络，SVM代表第二分类网络， ϕ 代表第一神经网络的特征提取网络，应理解，此处举例仅为方便理解本方案，不用于限定本方案。

本申请实施例中，训练设备先根据 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的特征信息，生成第一查询样本归属于 M 个类别中每个类别的概率，再根据第一查询样本归属于 M 个类别中每个类别的概率确定第一查询样本归属于 N 个类别中每个类别的概率，从而保证在一个小样本任务中，训练设备在对一个第一查询样本进行分类的过程中，学习到了 M 个类别中每个类别的特征，以提高训练后的第一神经网络的输出结果的精度。

307、训练设备根据第一损失函数和第二损失函数，对第一神经网络进行迭代训练，直至满足预设条件，第一损失函数指示第一预测结果和第二标注结果之间的相似度，第二损失函数指示第一预测结果与第二预测结果之间的相似度，或者，第二损失函数指示第二预测结果和第二标注结果之间的相似度。

本申请实施例中，在一次训练过程中，训练设备在得到每个第一查询样本的第一预测结果和每个第一查询样本的第二预测结果之后，可以生成第一损失函数的函数值和第二损失函数的函数值，并利用第一损失函数的函数值和第二损失函数的函数值，进行反向求导，以更新第一神经网络的权重参数，从而完成对第一神经网络的一次训练。

其中，第一损失函数指示每个第一查询样本的第一预测结果（也即预测类别）和每个第一查询样本的第二标注结果（也即正确类别）之间的相似度。第二损失函数指示每个第一查询样本的第一预测结果与每个第一查询样本的第二预测结果之间的相似度，或者，第二损失函数指示每个第一查询样本的第二预测结果和每个第一查询样本的第二标注结果之间的相似度。

进一步地，第一损失函数具体可以为交叉熵损失函数、相对熵损失函数或其他类型的

损失函数，第二损失函数可以为交叉熵损失函数、相对熵损失函数或其他类型的损失函数等，此处不做穷举。

具体的，在一次训练过程中，训练设备在生成第一损失函数的函数值和第二损失函数的函数值之后，可以对第一损失函数的函数值和第二损失函数的函数值进行加权求和，以得到总的函数值，对总的函数值进行梯度求导，并反向更新第一神经网络包括的特征提取网络的权重参数。第一损失函数的函数值和第二损失函数的函数值的权重为一个超参数。第一损失函数和第二损失函数的公式的一个示例可以如下：

$$\sum_{(x_j^S, y_j^S) \in \mathcal{D}_{\text{train}}^S} \ell(f(\mathcal{D}_{\text{train}}^S)(x_j^S), y_j^S) + \lambda \ell_M(f(\mathcal{D}_{\text{train}}^S)(x_j^S), h^S(x_j^S)); \quad (3)$$

其中，公式（3）代表第一损失函数和第二损失函数的和， $\mathcal{D}_{\text{train}}^S$ 代表与一个小样本任务对应的第一样本子集合， x_j^S 代表第一样本子集合中的第一查询样本， y_j^S 代表第一查询样本的正确类别（也即第二标注结果）， $f(\mathcal{D}_{\text{train}}^S)$ 代表第一查询样本在 N 类中的预测类别（也即第一预测结果）， $\ell(f(\mathcal{D}_{\text{train}}^S)(x_j^S), y_j^S)$ 代表第一损失函数，指示第一预测结果和第二标注结果之间的相似度， $h^S(x_j^S)$ 代表通过第二分类网络得到的第一查询样本在 N 类中的预测类别（也即第二预测结果）， $\ell_M(f(\mathcal{D}_{\text{train}}^S)(x_j^S), h^S(x_j^S))$ 代表第二损失函数，指示第一预测结果和第二预测结果之间的相似度， λ 代表第二损失函数的权重，是一个超参数。

公式（3）的训练目标为越来越小，也即训练目标为拉近每个第一查询样本的第一预测结果和每个第一查询样本的第二标注结果之间的相似度，且拉近每个第一查询样本的第一预测结果与每个第一查询样本的第二预测结果之间的相似度。应理解，公式（3）仅为方便理解第一损失函数和第二损失函数所做的一个示例，第二损失函数还可以用于指示每个第一查询样本的第二预测结果和每个第一查询样本的第二标注结果之间的相似度。

在一种情况下，训练设备在基于与一个小样本任务对应的一个第一样本子集合，对第一神经网络执行了一次训练之后，可以重新进入步骤 303，以利用同一个第一样本子集合对第一神经网络迭代训练，直至满足第一预设条件后，重新进入步骤 302，以从样本集合中采样出与另一个小样本任务对应的另一个第一样本子集合，进而基于另一个第一样本子集合对第一神经网络进行迭代训练直至满足第一预设条件。第一预设条件可以为迭代此次满足预设次数，也可以为满足第一损失函数和第二损失函数的收敛条件。训练设备重复利用预设个数的小样本任务对第一神经网络进行迭代训练，预设个数的取值可以为 20000 个、30000 个或其他数量等。

在另一种情况下，训练设备在基于与一个小样本任务对应的一个第一样本子集合，对第一神经网络执行了一次训练之后，可以重新进入步骤 302，以从样本集合中采样出与另一个小样本任务对应的另一个第一样本子集合，进而基于另一个第一样本子集合对第一神经网络进行下一次训练，直至满足第二预设条件。第二预设条件为迭代次数达到预设次数，

预设次数可以为 20000 次、30000 次或其他数量等。

为更直观的理解本方案，请参阅图 8，图 8 为本申请实施例提供的神经网络的训练方法的一种流程示意图。训练设备将第一样本子集合输入至第一神经网络的特征提取网络中，以通过特征提取网络对每个第一支持图像和每个第一查询图像进行特征提取，得到每个第一支持图像的特征信息和每个第一查询图像的第一特征信息。训练设备根据每个第一支持图像的特征信息生成 4 个类中心，4 个类中心分别为指示类别为狗的所有支持图像的特征、类别为狼的所有支持图像的特征、类别为桌子的所有支持图像的特征和类别为沙发的所有支持图像的特征，训练设备将 4 个类中心和每个第一查询图像的第一特征信息输入至第一分类网络中，得到第一分类网络输出的每个第一查询图像的第一预测结果。训练设备将每个第一查询图像的第一特征信息、与 M 组支持样本对应的 M 个第二特征信息以及与 M 组支持样本对应的 M 个第一标注结果，输入至第二分类网络，以通过第二分类网络得到每个第一查询图像的第二预测结果。训练设备根据每个第一查询图像的第一预测结果和每个第一查询图像的第二标注结果，生成第一损失函数的函数值，第一损失函数指示第一预测结果和第二标注结果之间的相似度；根据每个第一查询图像的第一预测结果和每个第一查询图像的第二预测结果，生成第二损失函数的函数值，第二损失函数指示第一预测结果与第二预测结果之间的相似度。训练设备根据第一损失函数的函数值和第二损失函数的函数值更新第一神经网络的权重参数，以实现了对第一神经网络的一次训练，应理解，图 8 中的举例仅为方便理解本方案的一个示例，不用于限定本方案。

本申请实施例中，可选地，训练设备在利用多个小样本任务对第一神经网络进行迭代训练，直至满足预设条件后，可以得到与多个小样本任务一一对应的多个训练后的第一神经网络。训练设备可以利用测试样本集合，对与多个小样本任务一一对应的多个训练后的第一神经网络进行精度测试，以挑选一个精度最高的训练后的第一神经网络。

训练设备可以从与目标任务对应的样本集合中采样得到与目标任务对应的小样本任务，训练设备利用与目标任务对应的小样本任务对精度最高的训练后的第一神经网络再次进行迭代训练，直至满足第一预设条件，得到最终训练后的第一神经网络。前述迭代训练中每次训练的过程可以参阅步骤 302 至步骤 307 的描述，此处不再赘述。

本申请实施例中，由于最后得到的神经网络过多学习同一小样本任务中不同样本在特征维度的区别，而忽略了与多个不同的小样本任务对应的多个不同类别的样本在特征维度的区别，从而导致神经网络输出的预测结果的精度较低。为了解决前述问题，本方案的训练阶段中，在每次小样本任务中，还会根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的特征信息，生成与第一查询样本对应的第二预测结果，并增设第二损失函数指示第一预测结果与第二预测结果之间的相似度，或者，指示第二预测结果和正确结果之间的相似度。从而使神经网络在一个小样本任务的训练过程中不仅能够学习到一个小样本任务中支持样本的特征，而且能够学习到 M 组支持样本的特征，也即神经网络在一个小样本学习中不仅能够学习到一个小样本任务中不同样本在特征维度的区别，而且能够学习到与多个不同小样本任务对应的多个不同类别的样本在特征维度的区别，从而提高输出的预测结果的精度。

(2) 训练的为用于回归的神经网络

本申请实施例中，请参阅图 9，图 9 为本申请实施例提供的神经网络的训练方法的一种流程示意图，方法可以包括：

901、训练设备创建特征信息库。

5 本申请实施例中，与图 3 对应实施例中步骤 301 类似，在训练后的神经网络用于回归的场景中，训练设备在对神经网络进行训练之前，也会创建特征信息库。特征信息库又称为元存储库 (meta-repository)，特征信息库包括的 M 个队列与样本集合中包括的 M 组支持样本一一对应。 M 个队列中的每个队列均用于存储与一组支持样本对应的特征信息和与一组支持样本对应的标注结果。

10 M 个队列中的每个队列还包括整个队列的标识信息和队列中每个支持样本的标注结果。与图 3 对应实施例的区别在于，第一，在图 3 对应实施例中， M 个队列中的一整个队列可以对应一个标注结果，一个队列中多个支持样本可以共享一个标注结果（也即正确类别或标签）；在图 9 对应实施例中， M 个队列中的任一个队列包括每个支持样本的标注结果，也即一个队列中标注结果的数量与支持样本的数量相同，每个支持样本的标注结果为每个支持样本的正确回归结果。作为示例，例如需要训练的第一神经网络的功能为电池充电进行个性化管理，支持样本的标注结果可以为充电时长为 6 小时、充电时长为 3 小时、充电时长为 9 小时等，此处举例仅为方便理解本方案，不用于限定本方案。

15 第二，在图 3 对应实施例中， M 个队列中一个队列的标识信息与一整个队列的标注结果可以为同一个信息。在图 9 对应实施例中， M 个队列中一个队列的标识信息和一个队列中每个支持样本的标注结果为不同的信息。本应用场景下， M 个队列中一个队列的标识信息具体可以表现为字符编码，该字符编码可以为随机生成的，也可以按照一定规则生成的。

902、训练设备从样本集合中获取第一样本子集合，第一样本子集合包括第一支持样本和第一查询样本。

20 本申请实施例中，步骤 902 的具体实现方式，可以参阅图 3 对应实施例中步骤 302 的描述，此处不做赘述。与图 3 对应实施例的区别在于，在图 3 对应实施例中，一个支持样本或一个查询样本可以为图像、文本或语音，一个支持样本或一个查询样本的标注结果为样本的正确类别。在图 9 对应实施例中，支持样本和查询样本具体表现为待回归的离散数据。

30 作为示例，例如神经网络用于对电池充电进行个性化管理，一个支持样本或一个查询样本中包括的可以为某个用户的手机型号、电池型号、充电地点和充电时刻；对应的，一个支持样本或一个查询样本的标注结果为正确的充电时长。由于手机可以控制充电速度，在充相同电量的前提下，预定的充电时长越长，则手机可以控制充电速度放缓；预定的充电时长越短，则手机可以控制充电速度加快，以保护手机电池，提高手机电池寿命。进一步地，例如充电地点是家、充电时刻是晚上 11 点，则一般情况下，手机与充电器的连接时长会比较长；例如充电地点是办公室，充电时刻是早上 10 点，则手机与充电器的连接时长会比较短等，不同用户的个人习惯、手机型号、电池型号等也可以影响第一神经网络的输出。应理解，此处举例仅为方便理解本应用场景，训练后的第一神经网络还可以应用于其

35

他的回归场景中，此处不做限定。

进一步地，在本应用场景下，**M** 组的分组依据为以下中的一项或多项：与支持样本对应的用户、地点、时间或其他元素等。本实施例中，第一样本子集合中包括的至少一个第一支持样本来自于所述 **M** 组的同一组，所述第一样本子集合中包括的至少一个第一查询样本来自于所述 **M** 组的同一组。本申请实施例中，提供了在回归这一应用场景下，**M** 组支持样本的多种分组依据，提高了本方案的实现灵活性。

具体的，在一种情况下，以分组依据为用户为例。在一种实现方式中，训练设备在得到样本集合后，针对样本集合中包括的支持样本集合，训练设备可以直接将每个用户的支持样本确定为一组支持样本，从而将 **M** 个用户的支持样本确定为 **M** 组支持样本。在另一种实现方式中，训练设备可以根据用户对支持样本集合进行分类，得到多个用户中每个用户的支持样本，进而根据每个用户的支持样本得到每个用户的特征信息，并根据多个用户中每个用户的特征信息执行聚类操作，以将样本集合包括的所有支持样本分为 **M** 组支持样本。将样本集合包括的所有查询样本分为 **M** 组查询样本的过程与前述描述类似，此处不做赘述。

在另一种情况下，以分组依据为地点为例。在一种实现方式中，训练设备在得到样本集合后，针对样本集合中包括的支持样本集合，训练设备可以直接将每个地点的支持样本确定为一组支持样本，从而将 **M** 个地点的支持样本确定为 **M** 组支持样本。在另一种实现方式中，训练设备可以根据地点对支持样本集合进行分类，得到与多个地点中每个地点对应的支持样本，进而根据与每个地点对应的支持样本得到每个地点的特征信息，并根据多个地点中每个地点的特征信息执行聚类操作，以将样本集合包括的所有支持样本分为 **M** 组支持样本。将样本集合包括的所有查询样本分为 **M** 组查询样本的过程与前述描述类似，此处不做赘述。

对于分类依据为时间的具体实现方式可以参阅分类依据为用户或地点的具体实现方式的描述，此处不做赘述。

903、训练设备将第一样本子集合输入至第一神经网络，获取第一神经网络生成第一查询样本的第一特征信息和第一支持样本的特征信息，并获取第一神经网络输出的与第一查询样本对应的第一预测结果，第一神经网络用于对样本进行回归。

本申请实施例中，步骤 903 的具体实现方式，可以参阅图 3 对应实施例中步骤 303 的描述，此处不做赘述。其中，第一神经网络为对样本进行回归处理的神经网络；具体的，第一神经网络可以为全连接神经网络。第一预测结果具体表现为预测回归结果。

904、训练设备将第一支持样本的特征信息和第一支持样本的标注结果添加至特征信息库的第一队列中。

本申请实施例中，与图 3 对应实施例中步骤 304 类似，针对多个第一支持样本中的任意一个第一支持样本，训练设备在通过第一神经网络生成一个第一支持样本的特征信息之后，会将一个第一支持样本的特征信息添加至第一队列中。与图 3 对应实施例中步骤 304 的区别在于，步骤 904 中，训练设备还会将该一个第一支持样本的标注结果添加至第一队列中。训练设备确定一个第一支持样本和一个第一队列之间对应关系的方式，可以参阅步骤 304

中的描述，此处不做赘述。

905、训练设备根据特征信息库，获取与 M 组支持样本对应的第二特征信息和与 M 组支持样本对应的第一标注结果。

5 本申请实施例中，与图 3 对应实施例中步骤 305 类似，训练设备通过进行取平均值处理，得到与 M 组支持样本中一组支持样本对应的第二特征信息，具体可参阅图 3 对应实施例中步骤 305 中的描述。

与图 3 对应实施例中步骤 305 的区别在于，步骤 905 中，训练设备还会生成与 M 组支持样本对应的 M 个第一标注结果。具体的，针对与 M 组支持样本中任一组支持样本对应的一个第一标注结果。训练设备从特征信息库包括的 M 个队列中获取与第二队列包括的所有支持样本对应的多个标注结果，对与第二队列包括的所有支持样本对应的多个标注结果进行取平均值处理，得到与第二队列对应的一个第一标注结果（也即得到与 M 组支持样本中一组支持样本对应的一个第一标注结果）。

906、训练设备根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的第一特征信息，生成与第一查询样本对应的第二预测结果。

15 本申请实施例中，与图 3 对应实施例中步骤 306 类似，针对第一样本子集合包括的一个或多个第一查询样本中的任一个第一查询样本。训练设备在获取到与 M 组支持样本对应的 M 个第二特征信息、与 M 组支持样本对应的 M 个第一标注结果之后，根据 M 个第二特征信息和一个第一查询样本的第一特征信息，执行相似度计算操作，得到第一相似度信息，第一相似度信息指示一个第一查询样本的第一特征信息与 M 个第二特征信息之间的相似度，也即指示第一查询样本与 M 组支持样本中每组支持样本之间的相似度。训练设备根据第一相似度信息和与 M 组支持样本对应的 M 个第一标注结果，生成第二预测结果。

具体的，与图 3 对应实施例中步骤 306 不同的是，本应用场景下，训练设备根据该一个第一查询样本与 M 组支持样本中每组支持样本之间的相似度，生成 M 个第一标注结果中每个第一标注结果的权重，并对 M 个第一标注结果进行加权求和，以生成该一个第一查询结果的第二预测结果。进一步地，训练设备可以将第一查询样本与一组支持样本之间的相似度直接确定为与该组支持样本对应的第一标注结果的权重；也可以将第一查询样本与一组支持样本之间的相似度进行二次处理，得到与该组支持样本对应的第一标注结果的权重等，此处不做限定。

30 作为示例，例如 M 个第一标注结果中 M 的取值为 12，第一相似度信息指示一个第一查询样本与 M 组支持样本中每组支持样本的相似度分别为 (0.01, 0.8, 0.01, 0.01, 0.02, 0.05, 0.01, 0.01, 0.01, 0.02, 0.03, 0.02)，则与 12 组支持样本中每组支持样本对应的每个第一标注结果的权重分别为 (0.01, 0.8, 0.01, 0.01, 0.02, 0.05, 0.01, 0.01, 0.01, 0.02, 0.03, 0.02)，应理解，此处举例仅为方便理解本方案，不用于限定本方案。

35 907、训练设备根据第一损失函数和第二损失函数，对第一神经网络进行迭代训练，直至满足预设条件，第一损失函数指示第一预测结果和第二标注结果之间的相似度，第二损失函数指示第一预测结果与第二预测结果之间的相似度，或者，第二损失函数指示第二预测结果和第二标注结果之间的相似度。

本申请实施例中，步骤 907 的具体实现方式，可以参阅图 3 对应实施例中步骤 307 的描述，此处不做赘述。如下公开了第一损失函数和第二损失函数的公式的另一个示例：

$$\sum_{(x_j^s, y_j^s) \in \mathcal{D}_{\text{test}}^s} \ell(f(\mathcal{D}_{\text{train}}^s)(x_j^s), y_j^s) + \lambda \ell_M(h^s(x_j^s)(x_j^s), y_j^s); \quad (4)$$

其中，公式（4）代表第一损失函数和第二损失函数的和， $\mathcal{D}_{\text{train}}^s$ 代表与一个小样本任务对应的第一样本子集， x_j^s 代表第一样本子集中的第一查询样本， y_j^s 代表第一查询样本的正确回归结果（也即第二标注结果）， $f(\mathcal{D}_{\text{train}}^s)$ 代表第一查询样本的预测回归结果（也即第一预测结果）， $\ell(f(\mathcal{D}_{\text{train}}^s)(x_j^s), y_j^s)$ 代表第一损失函数，指示第一预测结果和第二标注结果之间的相似度， $h^s(x_j^s)$ 代表通过步骤 906 得到的第二预测结果， $\ell_M(f(\mathcal{D}_{\text{train}}^s)(x_j^s), y_j^s)$ 代表第二损失函数，指示第二预测结果和第二标注结果之间的相似度， λ 代表第二损失函数的权重，是一个超参数。

公式（4）的训练目标为越来越小，也即训练目标为拉近每个第一查询样本的第一预测结果和每个第一查询样本的第二标注结果之间的相似度，且拉近每个第一查询样本的第二预测结果与每个第一查询样本的第二标注结果之间的相似度。应理解，公式（4）仅为方便理解第一损失函数和第二损失函数所做的一个示例，不用于限定本方案。

本申请实施例中，训练后的第一神经网络不仅可以用于对样本进行分类，而且可以用于对样本进行回归，扩展了本方案的应用场景。

本申请实施例中，第一神经网络可以为全连接神经网络、卷积神经网络或残差神经网络，提高了本方案的实现灵活性。

二、推理阶段

本申请实施例中，推理阶段指的是上述图 2 中执行设备 210 利用训练后的目标模型/规则 201 对样本进行处理的过程。请参阅图 10，图 10 为本申请实施例提供的样本的处理方法的一种流程示意图，方法可以包括：

1001、执行设备获取第三支持样本和第三查询样本。

本申请实施例中，执行设备在得到训练后的第一神经网络后，在每次推理之前，需要先获取 N 类第三支持样本和 n 个第三查询样本，第三支持样本的含义与第一支持样本的含义类似，第三查询样本的含义与第一查询样本的含义类似，具体可参阅步骤 302 和步骤 902 中的描述，此处不做赘述。

1002、执行设备将第三支持样本和第三查询样本输入至第一神经网络，以通过第一神经网络输出的与第三查询样本对应的第三预测结果，第三预测结果为根据第三支持样本的特征信息和第三查询样本的特征信息之间的相似度生成的，第一神经网络是根据第一损失函数和第二损失函数训练得到的。

本申请实施例中，步骤 1002 的具体实现方式，可以参阅图 3 对应实施例中步骤 303 和图 9 对应实施例中步骤 903 的描述，此处不做赘述。

其中，第一损失函数指示第一神经网络输出的预测结果和正确结果之间的相似度，第二损失函数指示第一神经网络输出的预测结果和与第三查询样本对应的第二预测结果之间的相似度，或者，第二损失函数指示第二预测结果和正确结果之间的相似度，第二预测结果为根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的正确结果和第三查询样本的特征信息得到的，第三支持样本能够归属于第三组的支持样本，第三查询样本能够归属于第三组的查询样本，第三组为 M 组中的一个组。

第三预测结果与第二预测结果的含义类似，第二预测结果、第一损失函数和第二损失函数的含义均可以参阅图 3 和图 9 对应实施例中的描述，此处不做赘述。

本申请实施例中，由于最后得到的神经网络过多学习同一个小样本任务中不同样本在特征维度的区别，而忽略了与多个不同的小样本任务对应的多个不同类别的样本在特征维度的区别，从而导致神经网络输出的预测结果的精度较低。第一神经网络是根据第一损失函数和第二损失函数训练生成的，第二损失函数指示第一预测结果与第二预测结果之间的相似度，或者，指示第二预测结果和正确结果之间的相似度，第二预测结果是根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的特征信息生成的。训练后的神经网络不仅学习到一个小样本任务中不同样本在特征维度的区别，而且能够学习到与多个小样本任务对应的多个不同类别的样本在特征维度的区别，从而神经网络输出的结果的精度较高。

为了对有益效果有进一步了解，以下结合数据来展示本申请实施例所带来的效果。在公开图像集 MiniImageNet 上进行了验证，结果如下表 1 和表 2。

表 1

训练方法	5-way 1-shot		5-way 5-shot	
	卷积神经网络	残差神经网络	卷积神经网络	残差神经网络
Proto	49.42	-	68.2	-
FEAT	55.15	62.96	71.61	78.49
本申请实施例	55.37	67.18	72.27	82.59

其中，如表 1 所示，5-way 1-shot 指的是在一个小样本任务中包括 5 类第一支持样本，每类第一支持样本中有 1 个第一支持样本，5-way 1-shot 指的是在一个小样本任务中包括 5 类第一支持样本，每类第一支持样本中有 5 个第一支持样本。Proto 和 FEAT (few-shot embedding adaptation w/ transformer) 是目前两种元学习的训练方法，当小样本任务为 5-way 1-shot，且第一神经网络选择卷积神经网络时，采用 Proto 的方式进行训练得到的训练后第一神经网络的精度为 49.42，采用 FEAT 的方式进行训练得到的训练后第一神经网络的精度为 55.15，采用本申请实施例的方式进行训练得到的训练后第一神经网络的精度为 55.37。表 1 中的其他数据可类推理解，此处不做赘述，很明显，采用本申请实施例提供的训练方法，在小样本场景中，训练后的第一神经网络的精度有所提升。

表 2

5-way 30-shot	卷积神经网络	残差神经网络
Proto	77.01±0.14	87.64±0.40
FEAT	78.53±0.56	87.92±0.42
本申请实施例	80.26±0.50	88.20±0.41

其中，如表 2 所示，5-way 30-shot 指的是在一个小样本任务中包括 5 类第一支持样本，每类第一支持样本中有 30 个第一支持样本，当第一神经网络选取卷积神经网络时，采用 Proto 的方式进行训练得到的训练后第一神经网络的精度为 77.01±0.14，采用 FEAT 的方式进行训练得到的训练后第一神经网络的精度为 78.53±0.56，采用本申请实施例的方式进行训练得到的训练后第一神经网络的精度为 80.26±0.50。表 1 中的其他数据可类推理解，此处不做赘述，很明显，采用本申请实施例提供的训练方法，在中量样本场景中，训练后的第一神经网络的精度有所提升。

在图 1 至图 10 所对应的实施例的基础上，为了更好的实施本申请实施例的上述方案，下面还提供用于实施上述方案的相关设备。具体参阅图 11，图 11 为本申请实施例提供的神经网络的训练装置的一种结构示意图。神经网络的训练装置 1100 可以包括获取模块 1101、输入模块 1102、生成模块 1103 和训练模块 1104。其中，获取模块 1101，用于从样本集合中获取第一样本子集合，样本集合包括 M 组支持样本和 M 组查询样本，第一样本子集合包括第一支持样本和第一查询样本；输入模块 1102，用于将第一样本子集合输入至第一神经网络，第一神经网络生成第一查询样本的第一特征信息，并得到第一神经网络输出的与第一查询样本对应的第一预测结果，第一预测结果为根据第一支持样本的特征信息和第一特征信息之间的相似度生成的；生成模块 1103，用于根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的第一特征信息，生成与第一查询样本对应的第二预测结果，第一标注结果中的每一个均指示与 M 组支持样本中一组支持样本对应的正确结果；训练模块 1104，用于根据第一损失函数和第二损失函数，对第一神经网络进行迭代训练，直至满足预设条件。第一损失函数指示第一预测结果和第二标注结果之间的相似度，第二损失函数指示第一预测结果与第二预测结果之间的相似度，或者，第二损失函数指示第二预测结果和第二标注结果之间的相似度，第二标注结果为第一查询样本的正确结果。

在一种可能的设计中，请参阅图 12，图 12 为本申请实施例提供的神经网络的训练装置的一种结构示意图。生成模块 1103，还用于通过第一神经网络生成第一支持样本的特征信息，第一支持样本选自于第一组，第一组为 M 组中的一个组。神经网络的训练装置 1100 还包括：添加模块 1105，用于将第一支持样本的特征信息添加至特征信息库的第一队列中，特征信息库包括与 M 组支持样本一一对应的 M 个队列，第一队列为 M 个队列中与第一组对应的一个队列，用于存储与第一组的支持样本对应的特征信息；获取模块 1101，还用于根据特征信息库，获取与 M 组支持样本对应的第二特征信息。

在一种可能的设计中，第二队列为 M 个队列中任一队列，第二队列包括至少一个支持样本的特征信息；获取模块 1101，具体用于对第二队列包括的所有支持样本的特征信息进行

行取平均值处理，得到与第二队列对应的一个第二特征信息。

在一种可能的设计中，生成模块 1103，具体用于：根据第二特征信息和第一特征信息，执行相似度计算操作，得到第一相似度信息，第一相似度信息指示第一查询样本与 M 组支持样本中每组支持样本之间的相似度；根据第一相似度信息和第一标注结果，生成第二预测结果。

在一种可能的设计中，第一神经网络为对样本进行分类的神经网络，第一样本子集合包括 N 类第一支持样本，第一预测结果指示第一查询样本在 N 类中的预测类别；生成模块 1103，具体用于：根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的特征信息，生成第一分类类别，第一分类类别指示第一查询样本在 M 类中的预测类别；从第一分类类别中获取第二预测结果，第二预测结果指示第一查询样本在 N 类中的预测类别。

在一种可能的设计中，第一神经网络为对样本进行回归处理的神经网络。

在一种可能的设计中，M 组的分组依据为以下中的一项或多项：与支持样本对应的用户、地点或时间。

在一种可能的设计中，第一样本子集合中包括的至少一个第一支持样本来自于 M 组的同一组，第一样本子集合中包括的至少一个第一查询样本来自于 M 组的同一组。

在一种可能的设计中，第一神经网络为以下中的任一种神经网络：全连接神经网络、卷积神经网络或残差神经网络。

需要说明的是，神经网络的训练装置 1100 中各模块/单元之间的信息交互、执行过程等内容，与本申请中图 3 至图 9 对应的各个方法实施例基于同一构思，具体内容可参见本申请前述所示的方法实施例中的叙述，此处不再赘述。

本申请实施例还提供一种样本处理装置，具体参阅图 13，图 13 为本申请实施例提供的样本处理装置的一种结构示意图。样本处理装置 1300 可以包括获取模块 1301 和输入模块 1302。其中，获取模块 1301，用于获取第三支持样本和第三查询样本；输入模块 1302，用于将第三支持样本和第三查询样本输入至第一神经网络，以通过第一神经网络输出的与第三查询样本对应的第三预测结果，第三预测结果为根据第三支持样本的特征信息和第三查询样本的特征信息之间的相似度生成的。第一神经网络是根据第一损失函数和第二损失函数训练得到的，第一损失函数指示第一神经网络输出的预测结果和正确结果之间的相似度，第二损失函数指示第一神经网络输出的预测结果和与第三查询样本对应的第二预测结果之间的相似度，或者，第二损失函数指示第二预测结果和正确结果之间的相似度，第二预测结果为根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的正确结果和第三查询样本的特征信息得到的，第三支持样本能够归属于第三组的支持样本，第三查询样本能够归属于第三组的查询样本，第三组为 M 组中的一个组。

需要说明的是，样本处理装置 1300 中各模块/单元之间的信息交互、执行过程等内容，与本申请中图 10 对应的各个方法实施例基于同一构思，具体内容可参见本申请前述所示的方法实施例中的叙述，此处不再赘述。

本申请实施例还提供一种训练设备，请参阅图 14，图 14 为本申请实施例提供的训练

设备的一种结构示意图。训练设备 1400 上可以部署有图 11 和图 12 对应实施例中所描述的神经网络的训练装置 1100，用于实现图 3 至图 9 对应实施例训练设备的功能。具体的，训练设备 1400 由一个或多个服务器实现，训练设备 1400 可因配置或性能不同而产生比较大的差异，可以包括一个或一个以上中央处理器（central processing units, CPU）1422（例如，一个或一个以上处理器）和存储器 1432，一个或一个以上存储应用程序 1442 或数据 1444 的存储介质 1430（例如一个或一个以上海量存储设备）。其中，存储器 1432 和存储介质 1430 可以是短暂存储或持久存储。存储在存储介质 1430 的程序可以包括一个或一个以上模块（图示没标出），每个模块可以包括对训练设备中的一系列指令操作。更进一步地，中央处理器 1422 可以设置为与存储介质 1430 通信，在训练设备 1400 上执行存储介质 1430 中的一系列指令操作。

训练设备 1400 还可以包括一个或一个以上电源 1426，一个或一个以上有线或无线网络接口 1450，一个或一个以上输入输出接口 1458，和/或，一个或一个以上操作系统 1441，例如 Windows Server™，Mac OS X™，Unix™，Linux™，FreeBSD™ 等等。

本申请实施例中，在中央处理器 1422 用于执行图 3 至图 9 对应实施例中的训练设备执行的神经网络的训练方法的情况下，具体的，中央处理器 1422 用于从样本集合中获取第一样本子集合，样本集合包括 M 组支持样本和 M 组查询样本，第一样本子集合包括第一支持样本和第一查询样本；将第一样本子集合输入至第一神经网络，第一神经网络生成第一查询样本的第一特征信息，并得到第一神经网络输出的与第一查询样本对应的第一预测结果，第一预测结果为根据第一支持样本的特征信息和第一特征信息之间的相似度生成的；根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的第一标注结果和第一查询样本的第一特征信息，生成与第一查询样本对应的第二预测结果，第一标注结果中的每一个均指示与 M 组支持样本中一组支持样本对应的正确结果；根据第一损失函数和第二损失函数，对第一神经网络进行迭代训练，直至满足预设条件；其中，第一损失函数指示第一预测结果和第二标注结果之间的相似度，第二损失函数指示第一预测结果与第二预测结果之间的相似度，或者，第二损失函数指示第二预测结果和第二标注结果之间的相似度，第二标注结果为第一查询样本的正确结果。

需要说明的是，对于中央处理器 1422 执行图 3 至图 9 对应实施例中神经网络的训练方法的具体实现方式以及带来的有益效果，均可以参考图 3 至图 9 对应的各个方法实施例中的叙述，此处不再一一赘述。

本申请实施例还提供一种执行设备，请参阅图 15，图 15 为本申请实施例提供的执行设备的一种结构示意图。其中，执行设备 1500 上可以部署有图 13 对应实施例中所描述的样本处理装置 1300，用于实现图 10 对应实施例中执行设备的功能。具体的，执行设备 1500 包括：接收器 1501、发射器 1502、处理器 1503 和存储器 1504（其中执行设备 1500 中的处理器 1503 的数量可以一个或多个，图 15 中以一个处理器为例），其中，处理器 1503 可以包括应用处理器 15031 和通信处理器 15032。在本申请的一些实施例中，接收器 1501、发射器 1502、处理器 1503 和存储器 1504 可通过总线或其它方式连接。

存储器 1504 可以包括只读存储器和随机存取存储器，并向处理器 1503 提供指令和数

据。存储器 1504 的一部分还可以包括非易失性随机存取存储器 (non-volatile random access memory, NVRAM)。存储器 1504 存储有处理器和操作指令、可执行模块或者数据结构, 或者它们的子集, 或者它们的扩展集, 其中, 操作指令可包括各种操作指令, 用于实现各种操作。

5 处理器 1503 控制执行设备的操作。具体的应用中, 执行设备的各个组件通过总线系统耦合在一起, 其中总线系统除包括数据总线之外, 还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见, 在图中将各种总线都称为总线系统。

 上述本申请实施例揭示的方法可以应用于处理器 1503 中, 或者由处理器 1503 实现。处理器 1503 可以是一种集成电路芯片, 具有信号的处理能力。在实现过程中, 上述方法的
10 各步骤可以通过处理器 1503 中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器 1503 可以是通用处理器、数字信号处理器 (digital signal processing, DSP)、微处理器或微控制器, 还可进一步包括专用集成电路 (application specific integrated circuit, ASIC)、现场可编程门阵列 (field-programmable gate array, FPGA) 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。该处理器 1503 可以实现或者执行本申请实施例
15 中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本申请实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成, 或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器, 闪存、只读存储器, 可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器 1504, 处理器 1503 读取存储器 1504
20 中的信息, 结合其硬件完成上述方法的步骤。

 接收器 1501 可用于接收输入的数字或字符信息, 以及产生与执行设备的相关设置以及功能控制有关的信号输入。发射器 1502 可用于通过第一接口输出数字或字符信息; 发射器 1502 还可用于通过第一接口向磁盘组发送指令, 以修改磁盘组中的数据; 发射器 1502 还可以包括显示屏等显示设备。

25 本申请实施例中, 应用处理器 15031, 用于执行图 10 对应实施例中执行设备的功能。需要说明的是, 对于应用处理器 15031 执行图 10 对应实施例中执行设备的功能的具体实现方式以及带来的有益效果, 均可以参考图 10 对应的各个方法实施例中的叙述, 此处不再赘述。

 本申请实施例中还提供一种计算机可读存储介质, 该计算机可读存储介质中存储有程序, 当其在计算机上运行时, 使得计算机执行如前述图 3 至图 9 所示实施例描述的方法中
30 训练设备所执行的步骤; 或者, 使得计算机执行如前述图 10 所示实施例描述的方法中执行设备所执行的步骤。

 本申请实施例中还提供一种包括计算机程序产品, 当其在计算机上运行时, 使得计算机执行如前述图 3 至图 9 所示实施例描述的方法中训练设备所执行的步骤, 或者, 使得计
35 算机执行如前述图 10 所示实施例描述的方法中执行设备所执行的步骤。

 本申请实施例中还提供一种电路系统, 所述电路系统包括处理电路, 所述处理电路配置为执行如前述图 3 至图 9 所示实施例描述的方法中训练设备所执行的步骤, 或者, 所述

处理电路配置为执行如前述图 10 所示实施例描述的方法中执行设备所执行的步骤。

本申请实施例提供的神经网络的训练装置或执行设备具体可以为芯片，芯片包括：处理单元和通信单元，所述处理单元例如可以是处理器，所述通信单元例如可以是输入/输出接口、管脚或电路等。该处理单元可执行存储单元存储的计算机执行指令，以使训练设备内的芯片执行上述图 3 至图 9 所示实施例描述的神经网络的训练方法，或者，以使训练设备内的芯片执行上述图 10 所示实施例描述的样本的处理方法。可选地，所述存储单元为所述芯片内的存储单元，如寄存器、缓存等，所述存储单元还可以是所述无线接入设备端内的位于所述芯片外部的存储单元，如只读存储器（read-only memory, ROM）或可存储静态信息和指令的其他类型的静态存储设备，随机存取存储器（random access memory, RAM）等。

具体的，请参阅图 16，图 16 为本申请实施例提供的芯片的一种结构示意图，所述芯片可以表现为神经网络处理器 NPU 160，NPU 160 作为协处理器挂载到主 CPU (Host CPU) 上，由 Host CPU 分配任务。NPU 的核心部分为运算电路 1603，通过控制器 1604 控制运算电路 1603 提取存储器中的矩阵数据并进行乘法运算。

在一些实现中，运算电路 1603 内部包括多个处理单元（Process Engine, PE）。在一些实现中，运算电路 1603 是二维脉动阵列。运算电路 1603 还可以是一维脉动阵列或者能够执行例如乘法和加法这样的数学运算的其它电子线路。在一些实现中，运算电路 1603 是通用的矩阵处理器。

举例来说，假设有输入矩阵 A，权重矩阵 B，输出矩阵 C。运算电路从权重存储器 1602 中取矩阵 B 相应的数据，并缓存在运算电路中每一个 PE 上。运算电路从输入存储器 1601 中取矩阵 A 数据与矩阵 B 进行矩阵运算，得到的矩阵的部分结果或最终结果，保存在累加器（accumulator）1608 中。

统一存储器 1606 用于存放输入数据以及输出数据。权重数据直接通过存储单元访问控制器（Direct Memory Access Controller, DMAC）1605，DMAC 被搬运到权重存储器 1602 中。输入数据也通过 DMAC 被搬运到统一存储器 1606 中。

BIU 为 Bus Interface Unit 即，总线接口单元 1610，用于 AXI 总线与 DMAC 和取指存储器（Instruction Fetch Buffer, IFB）1609 的交互。

总线接口单元 1610（Bus Interface Unit，简称 BIU），用于取指存储器 1609 从外部存储器获取指令，还用于存储单元访问控制器 1605 从外部存储器获取输入矩阵 A 或者权重矩阵 B 的原数据。

DMAC 主要用于将外部存储器 DDR 中的输入数据搬运到统一存储器 1606 或将权重数据搬运到权重存储器 1602 中或将输入数据数据搬运到输入存储器 1601 中。

向量计算单元 1607 包括多个运算处理单元，在需要的情况下，对运算电路的输出做进一步处理，如向量乘，向量加，指数运算，对数运算，大小比较等等。主要用于神经网络中非卷积/全连接层网络计算，如 Batch Normalization（批归一化），像素级求和，对特征平面进行上采样等。

在一些实现中，向量计算单元 1607 能将经处理的输出的向量存储到统一存储器 1606。

例如，向量计算单元 1607 可以将线性函数和/或非线性函数应用到运算电路 1603 的输出，例如对卷积层提取的特征平面进行线性插值，再例如累加值的向量，用以生成激活值。在一些实现中，向量计算单元 1607 生成归一化的值、像素级求和的值，或二者均有。在一些实现中，处理过的输出的向量能够用作到运算电路 1603 的激活输入，例如用于在神经网络中的后续层中的使用。

控制器 1604 连接的取指存储器（instruction fetch buffer）1609，用于存储控制器 1604 使用的指令；统一存储器 1606，输入存储器 1601，权重存储器 1602 以及取指存储器 1609 均为 On-Chip 存储器。外部存储器私有于该 NPU 硬件架构。

其中，循环神经网络中各层的运算可以由运算电路 1603 或向量计算单元 1607 执行。

其中，上述任一处提到的处理器，可以是一个通用中央处理器，微处理器，ASIC，或一个或多个用于控制上述第一方面方法的程序执行的集成电路。

另外需说明的是，以上所描述的装置实施例仅仅是示意性的，其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。另外，本申请提供的装置实施例附图中，模块之间的连接关系表示它们之间具有通信连接，具体可以实现为一条或多条通信总线或信号线。

通过以上的实施方式的描述，所属领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件的方式来实现，当然也可以通过专用硬件包括专用集成电路、专用 CLU、专用存储器、专用元器件等来实现。一般情况下，凡由计算机程序完成的功能都可以很容易地用相应的硬件来实现，而且，用来实现同一功能的具体硬件结构也可以是多种多样的，例如模拟电路、数字电路或专用电路等。但是，对本申请而言更多情况下软件程序实现是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在可读取的存储介质中，如计算机的软盘、U 盘、移动硬盘、ROM、RAM、磁碟或者光盘等，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本申请各个实施例所述的方法。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。

所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线（DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存储的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数

据存储设备。所述可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质（例如，DVD）、或者半导体介质（例如固态硬盘（Solid State Disk，SSD））等。

权 利 要 求

1.一种神经网络的训练方法，其特征在于，所述方法包括：

从样本集合中获取第一样本子集合，所述样本集合包括 M 组支持样本，所述第一样本子集合包括第一支持样本和第一查询样本；

5 将所述第一样本子集合输入至第一神经网络，所述第一神经网络生成所述第一查询样本的第一特征信息，并得到所述第一神经网络输出的与所述第一查询样本对应的第一预测结果，所述第一预测结果为根据所述第一支持样本的特征信息和所述第一特征信息之间的相似度生成的；

10 根据与所述 M 组支持样本对应的第二特征信息、与所述 M 组支持样本对应的第一标注结果和所述第一查询样本的第一特征信息，生成与所述第一查询样本对应的第二预测结果，所述第一标注结果中的每一个均指示与所述 M 组支持样本中一组支持样本对应的正确结果；

根据第一损失函数和第二损失函数，对所述第一神经网络进行训练，直至满足预设条件；

15 其中，所述第一损失函数指示所述第一预测结果和第二标注结果之间的相似度，所述第二损失函数指示所述第一预测结果与所述第二预测结果之间的相似度，或者，所述第二损失函数指示所述第二预测结果和所述第二标注结果之间的相似度，所述第二标注结果为所述第一查询样本的正确结果。

2.根据权利要求 1 所述的方法，其特征在于，所述将所述第一样本子集合输入至第一神经网络之后，所述方法还包括：

通过所述第一神经网络生成所述第一支持样本的特征信息，所述第一支持样本选自于第一组，所述第一组为所述 M 组中的一个组；

25 将所述第一支持样本的特征信息添加至特征信息库的第一队列中，所述特征信息库包括与所述 M 组支持样本一一对应的 M 个队列，所述第一队列为所述 M 个队列中与所述第一组对应的一个队列，用于存储与所述第一组的支持样本对应的特征信息；

所述根据与所述 M 组支持样本对应的第二特征信息、与所述 M 组支持样本对应的第一标注结果和所述第一特征信息，生成与所述第一查询样本对应的第二预测结果之前，所述方法还包括：

根据所述特征信息库，获取与所述 M 组支持样本对应的第二特征信息。

30 3.根据权利要求 2 所述的方法，其特征在于，第二队列为所述 M 个队列中任一队列，所述第二队列包括至少一个支持样本的特征信息，所述根据所述特征信息库，获取与所述 M 组支持样本对应的第二特征信息，包括：

对所述第二队列包括的所有支持样本的特征信息进行取平均值处理，得到与所述第二队列对应的一个第二特征信息。

35 4.根据权利要求 1 所述的方法，其特征在于，所述根据与所述 M 组支持样本对应的第二特征信息、与所述 M 组支持样本对应的第一标注结果和所述第一特征信息，生成与所述第一查询样本对应的第二预测结果，包括：

根据所述第二特征信息和所述第一特征信息，执行相似度计算操作，得到第一相似度信息，所述第一相似度信息指示所述第一查询样本与所述M组支持样本中每组支持样本之间的相似度；

根据所述第一相似度信息和所述第一标注结果，生成所述第二预测结果。

- 5 5.根据权利要求1至4任一项所述的方法，其特征在于，所述第一神经网络为对样本进行分类的神经网络，所述第一样本子集合包括N类第一支持样本，所述第一预测结果指示所述第一查询样本在所述N类中的预测类别；

10 所述根据与所述M组支持样本对应的第二特征信息、与所述M组支持样本对应的第一标注结果和所述第一查询样本的特征信息，生成与所述第一查询样本对应的第二预测结果，包括：

根据与所述M组支持样本对应的第二特征信息、与所述M组支持样本对应的第一标注结果和所述第一查询样本的特征信息，生成第一分类类别，所述第一分类类别指示所述第一查询样本在所述M类中的预测类别；

15 从所述第一分类类别中获取所述第二预测结果，所述第二预测结果指示所述第一查询样本在所述N类中的预测类别。

6.根据权利要求1至4任一项所述的方法，其特征在于，所述第一神经网络为对样本进行回归处理的神经网络。

7.根据权利要求6所述的方法，其特征在于，所述M组的分组依据为以下中的一项或多项：与支持样本对应的用户、地点或时间。

20 8.根据权利要求6所述的方法，其特征在于，所述第一样本子集合中包括的至少一个第一支持样本来自于所述M组的同一组，所述第一样本子集合中包括的至少一个第一查询样本来自于所述M组的同一组。

9.根据权利要求1至4任一项所述的方法，其特征在于，所述第一神经网络为以下中的任一种神经网络：全连接神经网络、卷积神经网络或残差神经网络。

25 10.一种样本的处理方法，其特征在于，所述方法包括：

获取第三支持样本和第三查询样本；

将所述第三支持样本和所述第三查询样本输入至第一神经网络，以通过所述第一神经网络输出的与所述第三查询样本对应的第三预测结果，所述第三预测结果为根据所述第三支持样本的特征信息和所述第三查询样本的特征信息之间的相似度生成的；

30 其中，所述第一神经网络是根据第一损失函数和第二损失函数训练得到的，所述第一损失函数指示所述第一神经网络输出的预测结果和正确结果之间的相似度，所述第二损失函数指示所述第一神经网络输出的预测结果和与所述第三查询样本对应的第二预测结果之间的相似度，或者，所述第二损失函数指示所述第二预测结果和正确结果之间的相似度，所述第二预测结果为根据与M组支持样本对应的第二特征信息、与M组支持样本对应的正确结果和所述第三查询样本的特征信息得到的，所述第三支持样本能够归属于第三组的支持样本，所述第三查询样本能够归属于第三组的查询样本，所述第三组为所述M组中的一个组。

35

11.一种神经网络的训练装置，其特征在于，所述装置包括：

获取模块，用于从样本集合中获取第一样本子集合，所述样本集合包括M组支持样本，所述第一样本子集合包括第一支持样本和第一查询样本；

5 输入模块，用于将所述第一样本子集合输入至第一神经网络，所述第一神经网络生成所述第一查询样本的第一特征信息，并得到所述第一神经网络输出的与所述第一查询样本对应的第一预测结果，所述第一预测结果为根据所述第一支持样本的特征信息和所述第一特征信息之间的相似度生成的；

10 生成模块，用于根据与所述M组支持样本对应的第二特征信息、与所述M组支持样本对应的第一标注结果和所述第一查询样本的第一特征信息，生成与所述第一查询样本对应的第二预测结果，所述第一标注结果中的每一个均指示与所述M组支持样本中一组支持样本对应的正确结果；

训练模块，用于根据第一损失函数和第二损失函数，对所述第一神经网络进行迭代训练，直至满足预设条件；

15 其中，所述第一损失函数指示所述第一预测结果和第二标注结果之间的相似度，所述第二损失函数指示所述第一预测结果与所述第二预测结果之间的相似度，或者，所述第二损失函数指示所述第二预测结果和所述第二标注结果之间的相似度，所述第二标注结果为所述第一查询样本的正确结果。

12.根据权利要求11所述的装置，其特征在于，

20 所述生成模块，还用于通过所述第一神经网络生成所述第一支持样本的特征信息，所述第一支持样本选自于第一组，所述第一组为所述M组中的一个组；

所述装置还包括：添加模块，用于将所述第一支持样本的特征信息添加至特征信息库的第一队列中，所述特征信息库包括与所述M组支持样本一一对应的M个队列，所述第一队列为所述M个队列中与所述第一组对应的一个队列，用于存储与所述第一组的支持样本对应的特征信息；

25 所述获取模块，还用于根据所述特征信息库，获取与所述M组支持样本对应的第二特征信息。

13.根据权利要求12所述的装置，其特征在于，第二队列为所述M个队列中任一队列，所述第二队列包括至少一个支持样本的特征信息；

30 所述获取模块，具体用于对所述第二队列包括的所有支持样本的特征信息进行取平均值处理，得到与所述第二队列对应的一个第二特征信息。

14.根据权利要求11所述的装置，其特征在于，所述生成模块，具体用于：

根据所述第二特征信息和所述第一特征信息，执行相似度计算操作，得到第一相似度信息，所述第一相似度信息指示所述第一查询样本与所述M组支持样本中每组支持样本之间的相似度；

35 根据所述第一相似度信息和所述第一标注结果，生成所述第二预测结果。

15.根据权利要求11至14任一项所述的装置，其特征在于，所述第一神经网络为对样本进行分类的神经网络，所述第一样本子集合包括N类第一支持样本，所述第一预测结果

指示所述第一查询样本在所述 N 类中的预测类别；

所述生成模块，具体用于：

根据与所述 M 组支持样本对应的第二特征信息、与所述 M 组支持样本对应的第一标注结果和所述第一查询样本的特征信息，生成第一分类类别，所述第一分类类别指示所述
5 第一查询样本在所述 M 类中的预测类别；

从所述第一分类类别中获取所述第二预测结果，所述第二预测结果指示所述第一查询样本在所述 N 类中的预测类别。

16.根据权利要求 11 至 14 任一项所述的装置，其特征在于，所述第一神经网络为对样本进行回归处理的神经网络。

10 17.根据权利要求 16 所述的装置，其特征在于，所述 M 组的分组依据为以下中的一项或多项：与支持样本对应的用户、地点或时间。

18.根据权利要求 16 所述的装置，其特征在于，所述第一样本子集合中包括的至少一个第一支持样本来自于所述 M 组的同一组，所述第一样本子集合中包括的至少一个第一查询样本来自于所述 M 组的同一组。

15 19.根据权利要求 11 至 14 任一项所述的装置，其特征在于，所述第一神经网络为以下中的任一种神经网络：全连接神经网络、卷积神经网络或残差神经网络。

20、一种样本处理装置，其特征在于，所述装置包括：

获取模块，用于获取第三支持样本和第三查询样本；

20 输入模块，用于将所述第三支持样本和所述第三查询样本输入至第一神经网络，以通过所述第一神经网络输出的与所述第三查询样本对应的第三预测结果，所述第三预测结果为根据所述第三支持样本的特征信息和所述第三查询样本的特征信息之间的相似度生成的；

其中，所述第一神经网络是根据第一损失函数和第二损失函数训练得到的，所述第一损失函数指示所述第一神经网络输出的预测结果和正确结果之间的相似度，所述第二损失函数指示所述第一神经网络输出的预测结果和与第三查询样本对应的第二预测结果之间的相似度，或者，所述第二损失函数指示所述第二预测结果和正确结果之间的相似度，
25 所述第二预测结果为根据与 M 组支持样本对应的第二特征信息、与 M 组支持样本对应的正确结果和所述第三查询样本的特征信息得到的，所述第三支持样本能够归属于第三组的支持样本，所述第三查询样本能够归属于第三组的查询样本，所述第三组为所述 M 组中的一个组。

30 21.一种训练设备，其特征在于，包括处理器，所述处理器和存储器耦合，所述存储器存储有程序指令，当所述存储器存储的程序指令被所述处理器执行时实现权利要求 1 至 9 中任一项所述的方法。

22.一种执行设备，其特征在于，包括处理器，所述处理器和存储器耦合，所述存储器存储有程序指令，当所述存储器存储的程序指令被所述处理器执行时实现权利要求 10 所述
35 的方法。

23.一种计算机可读存储介质，其特征在于，包括程序，当其在计算机上运行时，使得计算机执行如权利要求 1 至 9 中任一项所述的方法，或者，使得计算机执行如权利要求 10

所述的方法。

24.一种电路系统，其特征在于，所述电路系统包括处理电路，所述处理电路配置为执行如权利要求 1 至 9 中任一项所述的方法，或者，所述处理电路配置为执行如权利要求 10 所述的方法。