

AI 模型的训练测评装置及方法

相关申请的交叉引用

本申请要求 2023 年 12 月 29 日递交的发明名称为“AI 模型的训练测评装置及方法”的中国专利申请 No. 202311869522.0 的优先权，该中国专利申请的全部内容通过引用并入本文中。

技术领域

本申请总体上涉及人工智能（AI）领域，并且更具体地涉及 AI 模型的训练测评装置及方法。

背景技术

近些年，随着人工智能技术的飞速发展，越来越多的 AI 模型被应用到各个领域，而一个 AI 模型的产生通常要经历模型训练和模型测评两个阶段。

模型训练是指通过大量的训练数据进行模型参数的训练、调优的过程；而模型测评则是指通过部分预先准备、设计好的测试数据对训练过程中产生的 AI 模型进行测试的过程。模型测评过程结束后通常会产生一些被测评模型在测试数据中的测试结果，例如图像、指标等。这些指标能够反应 AI 模型的效果，从而帮助算法人员挑选效果较好的模型来应用于生产中。

当前 AI 模型的训练和测评通常都是算法人员全手动完成的。例如，算法人员训练产生一个模型后会暂时中断训练，并手动取出模型运行测评流程对模型进行测评，在测评完成后根据模型测评的结果以决定是否继续发起训练；或者在训练全部完成后同时取出训练产生的多个模型，依次对每个模型运行测评流程，待所有测评流程运行完成后，再手动选择最优模型。

这种全手动方式的训练测评过程由于包含了太多的人工操作带来了测评效率低、人力成本高等问题。

发明内容

鉴于以上所述的问题，根据本申请实施例的一个方面，提供了一种用于 AI 模型的训练测评的装置，包括：存储器，用于存储程序指令；以及处理器，该处理器与存储器耦合，并且被配置为执行程序指令来：创建应用于容器化应用管理平台的流程任务模板，该流程任务模板包括：与 AI 模型的训练测评任务相关联的流程配置信息、以及训练测评任务中的各个步骤所对应的阶段标志；并且响应于接收到持续训练测评任务的启动命令，执行以下操作：查询流程任务模板，根据流程任务模板中的流程配置信息和阶段标志，对训练测评任务进行阶段拆分以构建训练阶段任务，根据训练阶段任务进行产物逃逸分析，以生成与训练阶段任务相关联的产物逃逸分析结果，基于与训练阶段任务相关联的产物逃逸结果，启动训练阶段任务，并且当训练阶段任务达到预定的训练阶段性节点时，在不中断训练阶段任务的情况下，启动一个或多个测评阶段任务的并发运行。

根据本申请实施例的另一方面，提供了用于 AI 模型的训练测评的方法，包括：创建应用于容器化应用管理平台的流程任务模板，该流程任务模板包括：与 AI 模型的训练测评任务相关联的流程配置信息、以及训练测评任务中的各个步骤所对应的阶段标志；并且响应于接收到持续训练测评任务的启动命令：查询流程任务模板，根据流程任务模板中的流程配置信息和阶段标志，对训练测评任务进行阶段拆分以构建训练阶段任务，根据训练阶段任务进行产物逃逸分析，以生成与训练阶段任务相关联的产物逃逸分析结果，基于与训练阶段任务相关联的产物逃逸结果，启动训练阶段任务，并且当训练阶段任务达到预定的训练阶段性节点时，在不中断训练阶段任务的情况下，启动一个或多个测评阶段任务的并发运行。

根据本申请实施例的又一方面，提供了一种机器可读介质，其上存储有程序指令，该程序指令在被处理器执行时使得处理器执行如上所述的用于 AI 模型的训练测评的方法。

根据本申请实施例的又一方面，提供了一种用于 AI 模型的训练测评的设备，包括用于执行如上所述的用于 AI 模型的训练测评的方法的装置。

根据本申请实施例的用于 AI 模型的训练测评的装置、方法、设备和机器可读介质，抽象并固化了 AI 模型的训练测评流程配置，实现了 AI 模型的持续训练测评，在训练测评过程中减少了人工的参与，解决了模型训练和模型测评无法并行进行的问题，从而可以提高 AI 模型开发的效率。

此外，根据本申请的一些实施例，还可以实现统一的 AI 模型单次/持续训练测评方法。在这些实施例中，平台可以响应于用户输入的单次训练测评启动命令或持续训练测评启动命令，相应地启动单次训练测评或持续训练测评程序。即，经过统一的训练测评配置后，用户既可以选择运行单次训练测评任务，也可以在不修改配置的情况下，选择运行持续训练测评任务。

附图说明

下面将通过参考附图来描述本申请示例性实施例的特征、优点和技术效果。

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中所需要使用的附图作简单的介绍，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 示出了根据本申请的一些实施例的 AI 模型的训练测评过程的示意性整体流程图；

图 2 示出了根据本申请的一些实施例的 AI 模型的训练测评过程中的流程配置的示例过程；

图 3 示出了根据本申请的一些实施例的 AI 模型的单次训练测评任务流程；

图 4 示出了根据本申请的一些实施例的 AI 模型的持续训练测评任务流程的整体流程图；

图 5 示出了根据本申请的一些实施例的 AI 模型的持续训练测评任务

流程中的阶段拆分过程的示意图；

图 6 示出了根据本申请的一些实施例的 AI 模型的持续训练测评任务流程中利用 sidecar 代理辅助实现产物文件的传递的过程的示意图；

图 7 示出了根据本申请的一些实施例的 AI 模型的持续训练测评任务流程中的非侵入的 Sidecar 代理容器注入的示意图；

图 8 示出了根据本申请的一些实施例的 AI 模型的持续训练测评任务流程中的多个模型测评任务并发运行的示意图；

图 9 示出了根据本申请的一些实施例的 AI 模型的持续训练测评任务流程的详细流程图；

图 10 示出了根据本申请的一些实施例的 AI 模型的训练测评装置的硬件结构示意图。

具体实施方式

下面将详细描述本申请的各个方面的特征和示例性实施例。在下面的详细描述中，提出了许多具体细节，以便提供对本申请的全面理解。但是，对于本领域技术人员来说很明显的是，本申请可以在不需要这些具体细节中的一些细节的情况下实施。下面对实施例的描述仅仅是为了通过示出本申请的示例来提供对本申请的更好的理解。本申请决不限于下面所提出的任何具体配置，而是在不脱离本申请的精神的前提下覆盖了元素、部件和算法的任何修改、替换和改进。在附图和下面的描述中，没有示出公知的结构和技术，以便避免对本申请造成不必要的模糊。

为解决在人工智能模型训练测评中使用全手动训练、测评工作流中带来的人工参与度高、测评效率低、人力成本高等问题，在本申请中提出了一种可以在容器化应用管理平台（例如，基于 Kubernetes 的平台）上实现的 AI 模型的持续训练测评过程，该过程抽象并固化了 AI 模型的训练测评流程配置，解决了模型训练和模型测评无法并行进行的问题。此外，根据本申请的实现 AI 模型的持续训练测评过程的平台还可以很方便地兼容 AI 模型的单次训练测评过程。因而，根据本申请的一些实施例，可以实现统一的 AI 模型单次或持续训练测评过程。下面将参考图 1 来描述这种统一

训练测评过程的整体流程，然后再参考图 2 至图 9 详细描述抽象并固化 AI 模型的训练测评流程配置的过程，以及单次训练测评过程和持续训练测评过程的具体流程，尤其是将着重描述持续训练测评过程的实现流程。

图 1 示出了根据本申请一些实施例的 AI 模型的训练测评过程的示意性整体流程图。如图 1 所示，AI 模型的训练测评的整体流程可以包括流程配置过程、启动训练过程、以及根据用户选择而启动的单次训练测评任务流程或持续训练测评任务流程。

总体而言，该训练测评过程可以在容器化应用管理平台上实现。首先，平台可以基于人工智能相关领域知识，构建 AI 模型的训练测评任务的任务步骤组成，该任务步骤组成包括完成 AI 模型的训练测评任务所需要的各个步骤。例如，对于图像检测领域，可以构建包括图像预处理步骤、模型训练步骤、模型测评步骤、后处理步骤等主要步骤的任务步骤组成。平台可以通过软件方法将这些任务步骤固化为流程任务模板，以便于后续基于该流程任务模板运行训练测评任务。然后，基于所固化的流程任务模板，平台可以响应于用户输入的训练测评启动命令，启动 AI 模型的单次训练测评任务或持续训练测评任务，从而可以实现 AI 模型的训练测评功能的自动化、并行化，并且可以提供一种统一的模型、报表的存储和展示方式。

在具体实现方式上，例如通过使用容器化应用管理引擎（例如，Kubernetes）、sidecar 机制、对象存储、流程引擎、软件开发工具包（SDK）等相关技术手段，将 AI 模型的训练测评任务流程进行模板化配置，同时通过阶段拆分、产物逃逸分析等机制，使得能够根据用户需要选择进行单次训练测评任务流程或持续训练测评任务流程。如果用户选择启动单次训练测评任务流程，则将运行单个包含 AI 模型的训练和测评阶段的完整流程，与全手动方式类似，但训练和测评流程中的各个步骤都将自动化顺序完成，无需过多手动干预。如果用户选择了持续训练测评任务流程，则将自动对流程任务进行阶段拆分和产物逃逸分析，基于阶段拆分和产物逃逸分析的结果启动训练阶段任务（在本文中也简称为训练任务），并且当训练阶段任务达到预定的训练阶段性节点时，在不中断所述训练阶

段任务的情况下，启动一个或多个测评阶段任务（在本文中也简称为测评任务）的并发运行，从而实现训练阶段任务和测评阶段任务的并行运行。

下面将结合图 2 至图 9 详细描述 AI 模型的训练测评流程配置的过程，以及单次训练测评过程和持续训练测评过程的具体流程，尤其是将参考图 4 至图 9 着重描述持续训练测评过程的具体实现流程。

图 2 示出了根据本申请的一些实施例的 AI 模型的训练测评过程中的流程配置的示例过程。根据本申请的实施例，流程配置是指通过软件方法，将 AI 相关领域的算法专家工作流中的相关运行代码、配置、运行环境等，对应至所抽象出的运行步骤（例如对于图像检测领域，可以包括图像预处理步骤、模型训练步骤、模型测评步骤、后处理步骤等）中，即，将 AI 相关领域的算法专家的工作流固化为流程任务模板的过程。

例如，根据本申请的实施例的技术方案可以在基于容器化应用管理引擎（例如，Kubernetes）的平台上实现，所有的训练、测评任务最终都是以容器（Pod）的方式被调度和运行，所以在具体实现时，AI 领域的算法专家工作流的相关配置可以被映射为例如容器启动命令、容器镜像、阶段产物、配置文件等参数，然后再在软件层面上根据这些参数生成流程任务模板，固化到 AI 模型的训练测评过程中。

具体而言，在流程配置阶段，可以对 AI 模型的训练测评流程进行映射和参数配置，以生成用于在容器化应用管理平台上运行的与 AI 模型的训练测评任务相关联的流程任务模板，该流程任务模板可以包括与 AI 模型的训练测评任务相关联的流程配置信息。如图 2 所示，例如，对于图像检测模型而言，流程配置信息可以包括与图像预处理步骤相关联的图像检测训练集、训练集预处理配置、图像检测测试集和测试集预处理配置；与模型训练步骤相关联的镜像配置、启动命令配置、资源配置、产物配置和输入文件配置；与模型测评步骤相关联的镜像配置、启动命令配置、资源配置、产物配置和输入文件配置；以及与后处理步骤相关联的模型测评报表分析和模型后处理配置。

此外，如上所述，根据本申请的一些实施例，当用户选择启动持续训练测评任务流程时，平台需要自动对流程任务进行阶段拆分和产物逃逸分

析，基于阶段拆分和产物逃逸分析的结果启动训练阶段任务，并且当训练阶段任务达到预定的训练阶段性节点时，在不中断所述训练阶段任务的情况下，启动一个或多个测评阶段任务的并发运行。

在这些实施例中，提出在创建流程任务模板时，基于与 AI 模型相关联的领域知识，构建 AI 模型的训练测评任务的任务步骤组成，该任务步骤组成包括完成 AI 模型的训练测评任务所需要的各个步骤；并且为任务步骤组成中的各个步骤分配该步骤所属的阶段，以得到与该步骤相对应的阶段标志。与各个步骤相对应的阶段标志可以被包括在流程任务模板中，来为后续阶段任务的构建提供基础。

例如，对于图像检测模型的训练测评任务而言，任务步骤组成通常包含图像预处理步骤、模型训练步骤、模型测评步骤和后处理步骤。可以基于图像检测领域知识对这些步骤进行阶段的划分，即，给每个步骤分配所属的阶段，为后续阶段任务的构建提供基础。例如，图像检测模型的训练测评任务可以被划分为两个阶段：模型训练阶段和模型测评阶段，为这两个阶段分别分配阶段标志 1 和 2，再为任务步骤组成中的各个步骤配设对应的阶段标志。例如，由于在训练阶段任务和测评阶段任务都需要图像预处理步骤，所以应为图像预处理步骤配设阶段标志 1 和 2，而模型训练步骤仅配设阶段标志 1，模型测评步骤仅配设阶段标志 2。

因此，在平台创建的流程任务模板中，不仅包括与 AI 模型的训练测评任务相关联的流程配置信息，还包括训练测评任务中的各个步骤所对应的阶段标志。这样，在平台需要启动某一阶段标志所对应的阶段任务时，可以根据流程任务模板中的各个步骤所对应的阶段标志，来构建出对应的要启动的阶段任务。例如，当需要启动模型训练阶段任务（对应于阶段标志 1）时，可以基于任务步骤组成中被配设的阶段标志包括阶段标志 1 的各个步骤来构建模型训练阶段任务，以便后续启动该训练阶段任务。

如图 1 所示，在完成 AI 模型的训练测评流程的流程任务模板的配置后，可以一键启动训练测评任务。在启动训练测评任务时，用户可以根据需要选择进行单次训练测评任务或持续训练测评任务。训练测评任务启动后，平台将会根据流程配置阶段中创建的流程任务模板，基于 Kubernetes

容器运行模型训练容器和模型测评容器，并监控其运行状态。

根据本申请的实施例，若在启动训练时用户选择启动单次训练测评任务，则平台将直接运行 AI 模型的单次训练测评任务，即，基于流程任务模板顺序执行 AI 模型的单次训练阶段任务和单次测评阶段任务，以完成该单次训练测评任务。图 3 示出了根据本申请的一些实施例的 AI 模型的单次训练测评任务流程。

如图 3 所示，以图像检测模型为例进行说明，图像检测模型的单次训练测评任务流程可以包括预处理阶段、训练阶段和测评阶段。具体而言，在图像检测模型的训练测评的预处理阶段，主要进行以下操作：自动化地根据图像检测领域专家配置的图像拉取条件，从对象存储中按需拉取训练、测试所需的图像数据，同时根据算法专家所配置的图像预处理算法，对拉取到的图像数据进行自动化的预处理，以生成预处理产物（图像数据）。在完成图像检测模型的训练测评任务所需的图像数据的拉取、图像数据的预处理后，将自动运行训练阶段容器。在训练阶段中，将运行图像检测模型的训练代码，使用预处理后的图像数据进行图像检测模型的训练。训练阶段将会产生一些产物文件，最主要的产物文件包含模型文件，软件层面会自动完成这些产物文件的存档等操作。在训练阶段完成后，将顺序执行图像检测模型的测评阶段，在训练阶段使用的基础模型文件和训练后产生的模型文件都会被自动传送至测评阶段容器，同时图像检测领域的算法专家所选择的相关图像测试集也会被自动传送至测评阶段。测评阶段容器通过运行测评代码，产生一些模型报表文件。对于图像检测模型，报表中通常会包含例如图像分割结果图、目标物检测图、目标检测正检率、目标检测误报率，以及目标检测漏报率等的指标，同样在该阶段，软件层面会自动完成这些产物文件的存档等操作。

根据本申请的一些实施例，若在启动训练时用户选择启动持续训练测评任务，则平台将自动对训练测评任务进行阶段拆分和产物逃逸分析，基于阶段拆分和产物逃逸分析的结果启动训练阶段任务，并且当训练阶段任务达到预定的训练阶段性节点时，可以在不中断训练阶段任务的情况下，启动一个或多个测评阶段任务的并发运行，以进行持续训练测评。图 4 示

出了根据本申请的一些实施例的 AI 模型的持续训练测评任务流程的整体流程图。

下面的表 1 示出了在根据本申请的示例实施例中，AI 模型的持续训练测评任务流程中所涉及的一些关键操作。表 1 中的序号对应于在图 4 中所示的这些操作的序号。

序号	操作	说明
1	调用 SDK 更新数据集产物	在持续训练过程中算法训练代码逻辑可以按需要使用 SDK 更新训练中使用到的数据集的预处理产物文件
2	SDK 调用 Sidecar 代理接口	SDK 通过接口与 Sidecar 代理进行交互，通知其更新数据集预处理产物文件或启动并发 AI 模型测评任务
3	Sidecar 更新数据集产物文件	Sidecar 代理实际进行本地数据集预处理产物文件的更新
4	训练产生产物	训练阶段经过模型迭代后生成如 AI 模型文件等相关训练产物
5	调用 SDK 通知模型产生和启动并行测评	算法训练代码使用 SDK 通告训练产物的产生并启动模型并行测评任务
6	Sidecar 代理获取训练产物	Sidecar 代理通过与训练容器共享文件系统，获取通告中的训练产物文件
7	存储产物至对象存储	Sidecar 代理根据产物文件的对象存储路径，完成本地产物文件到对象存储的上传过程
8	Sidecar 代理调用平台接口启动并	Sidecar 根据产物逃逸分析，整理模型测评所需的参数，调用平台接口并行启动模型测评

	行模型测评任务	任务
9	平台创建模型测评任务	平台服务根据接收到的参数创建对应模型的模型测评任务
10	模型测评任务获取输入文件	模型测评任务根据本任务的配置从对象存储中获取对应的输入文件，运行模型测评任务

表 1

如以上结合图 1 所描述的，持续训练测评任务流程将涉及阶段拆分、产物逃逸分析、产物存储和传递以及模型训练任务和模型测评任务的并行执行等多个方面。表 1 中所示的操作实际上就是与这些方面相关的操作。下面将结合图 5 至图 9 对与这些方面相关的操作进行更详细的描述。

图 5 示出了根据本申请的一些实施例的 AI 模型的持续训练测评任务流程中的阶段拆分过程的示意图。如图 5 所示，阶段拆分是指自动将 AI 模型的训练测评任务中的测评阶段任务从流程中拆分出来，即仅先启动训练阶段任务。具体而言，当平台接收到持续训练测评任务的启动命令时，平台将查询流程任务模板，根据流程任务模板中的流程配置信息和阶段标志，对训练测评任务进行阶段拆分，以构建 AI 模型的训练阶段任务。

如前所述，根据本申请的实施例，提出在创建流程任务模板时，基于与 AI 模型相关联的领域知识，构建 AI 模型的训练测评任务的任务步骤组成，并为任务步骤组成中的各个步骤分配该步骤所属的阶段，以得到与该步骤相对应的阶段标志。与各个步骤相对应的阶段标志可以被包括在流程任务模板中，来为后续阶段任务的构建提供基础。这样，在平台需要启动某一阶段标志所对应的阶段任务时，可以根据流程任务模板中的各个步骤所对应的阶段标志，来构建出对应的要启动的阶段任务。

基于流程任务模板中所配设的阶段标志，当需要启动对应于阶段标志 1 的训练阶段任务时，可以在阶段拆分过程中，基于任务步骤组成中被配设的阶段标志包括阶段标志 1 的各个步骤来构建训练阶段任务。

此外，根据本申请的实施例，平台在进行阶段拆分构建了训练阶段任务之后，还需要进行产物逃逸分析。具体而言，算法专家可以为训练测评

任务中所涉及的每个步骤配置相应的产物（例如，包括输入产物和输出产物），这些产物需要在多个步骤之间传递使用，而由于进行了阶段任务的拆分，原本应用于单一任务中的各个步骤之间的产物传递方式可能就不再适用了。因此，需要对产物跨阶段传递的情况进行分析和特别处理，在本文中也称为进行产物逃逸分析。

具体而言，在要启动某一阶段任务时，首先获取流程任务模板，流程任务模板中包含了算法专家配置的各个步骤的输入产物和输出产物，遍历这些步骤的输入产物和输出产物，分析得到这些产物在各个步骤间的传递关系（例如，产物 1 由步骤 1 产生，在步骤 2 中被作为输入）。在分析得到产物在各个步骤之间的传递关系之后，筛选出传递关系跨越训练测评任务中的不同阶段的产物（例如，产物需要在步骤 1 和步骤 2 之间传递，而步骤 1 属于阶段标志为 1 的训练阶段，步骤 2 属于阶段标志为 2 的测评阶段），这些产物就作为逃逸产物进行处理。当存在逃逸产物时，平台首先取消原本单一任务中对这些逃逸产物的原有处理逻辑，然后基于该阶段任务中的各个步骤对应的阶段标志，建立逃逸产物的传递处理逻辑，作为与该阶段任务相关联的产物逃逸分析结果。

需要注意的是，在启动训练阶段任务和测评阶段任务时，都可能出现产物跨阶段传递的情况，因而，都需要进行相应的产物逃逸分析。具体而言，在要启动训练阶段任务时进行的产物逃逸分析包括：基于流程任务模板，分析训练阶段任务中的各个步骤的产物在各个步骤之间的传递关系；筛选出传递关系跨越训练测评任务中的不同阶段的产物，作为逃逸产物进行处理；并且基于训练阶段任务中的各个步骤对应的阶段标志，建立逃逸产物的传递处理逻辑，作为与训练阶段任务相关联的产物逃逸分析结果。相应地，在要启动测评阶段任务时进行的产物逃逸分析包括：基于流程任务模板，分析测评阶段任务中的各个步骤的产物在各个步骤之间的传递关系；筛选出传递关系跨越训练测评任务中的不同阶段的产物，作为逃逸产物进行处理；并且基于测评阶段任务中的各个步骤对应的阶段标志，建立逃逸产物的传递处理逻辑，作为与测评阶段任务相关联的产物逃逸分析结果。

此外，关于产物逃逸分析结果的传递，根据本申请的实施例，提出可以通过以下方式进行产物逃逸分析结果的传递：为阶段任务配置关联的代理容器，利用该代理容器根据产物逃逸分析结果存储产物文件，并辅助实现产物文件在训练阶段任务和测评阶段任务之间的传递。

例如，可以将与训练阶段任务相关联的产物逃逸分析结果，通过环境变量的形式传递给 sidecar 代理容器，同时 sidecar 代理容器通过挂载同一块存储的方式与训练容器共享存储，即可以获取到训练阶段任务产生的产物；训练容器在训练任务达到阶段性成果后，通过专用的软件开发工具包 SDK 通知 sidecar 代理容器启动模型测评任务，sidecar 代理容器根据环境变量获取到的产物逃逸分析结果，对所有逃逸产物进行处理（如打包、存档等），最后通过平台服务接口启动模型测评任务，将本次逃逸产物的存档信息通过接口传递给平台服务。

图 6 示出了根据本申请的一些实施例的 AI 模型的持续训练测评任务流程中利用 sidecar 代理辅助实现产物文件的传递的过程的示意图。如图 6 所示，当产物传递跨越了不同的阶段任务时，则需要将该产物标记为逃逸产物，并且在后续发起独立的测评阶段任务时，可以由 sidecar 代理容器辅助进行产物文件的存储和传递。

根据本申请的一些实施例，为了实现在模型训练容器运行过程中可以自主地完成训练使用的数据集预处理产物更新、训练阶段产物（模型）的存储，以及并发模型测评任务的启动等操作，在技术实现上采用了非侵入的代理容器注入方式。该代理容器例如可以是基于用于容器化应用管理平台的 Sidecar 架构模式的 Sidecar 代理容器。

图 7 示出了根据本申请的一些实施例的 AI 模型的持续训练测评任务流程中的非侵入的 Sidecar 代理容器注入的示意图。由于训练阶段任务中发起存储训练阶段产物、更新数据集预处理产物等动作都包含复杂耗时的操作，而启动并发测评任务则需要调用平台服务接口，故直接由训练容器来实现这些操作是不合适的。为解决这一问题，在本申请的实施例中，提出采用 SDK 结合 Sidecar 代理容器的方式，由 Sidecar 代理容器完成复杂耗时的操作和调用平台服务接口，SDK 层面则仅封装简单的接口层。

如图 7 所示，训练阶段任务被配置有关联的 Sidecar 代理容器，该 Sidecar 代理容器可以通过专用的 SDK 与训练阶段任务和测评阶段任务交互，以进行产物的存储和产物在 Sidecar 代理容器与训练阶段任务和测评阶段任务之间的传递。具体而言，通过在 Kubernetes 上的训练任务所属的容器中自动注入自研的 Sidecar 代理容器，训练容器仅需与 Sidecar 代理容器通过本地网络进行通信（该通信也将由 python SDK 进行封装），同时 Sidecar 代理容器与训练容器可以共享部分文件系统，即，按需同步/异步地完成所需文件存储或传递操作，或调用远程平台服务接口来启动多个模型测评任务的并发运行。

图 8 示出了根据本申请的一些实施例的 AI 模型的持续训练测评任务流程中的多个模型测评任务并发运行的示意图。如图 8 所示，在根据本申请的实施例中，训练任务和测评任务可以互不影响地并行运行，同时多个测评任务也可以并发运行。例如，当训练阶段任务达到预定的训练阶段性节点时，可以在不中断训练阶段任务的情况下，启动一个或多个测评阶段任务的并发运行。在本申请的实施例中，预定的训练阶段性节点可以包括：训练阶段任务达到阶段性模型训练目标的节点、或者训练阶段任务达到预定义的训练迭代次数的节点。例如，当训练阶段任务达到预定的第一训练阶段性节点时，可以针对在第一训练阶段性节点处得到的训练模型启动第一测评阶段任务；当训练阶段任务达到预定的第二训练阶段性节点时，可以针对在第二训练阶段性节点处得到的训练模型启动第二测评阶段任务；以此类推。在启动多个测评阶段任务的并发运行的同时，可以继续继续进行训练阶段任务。另外，如表 1 中所示，可以利用 Sidecar 代理容器调用平台服务接口来启动多个模型测评任务的并发运行。

根据本申请的实施例，采用 SDK 结合 Sidecar 代理的方式来实现产物文件的存储和传递以及远程平台服务接口的调用，将具有如下优势：将这部分逻辑从 AI 模型的训练代码中剥离后，降低了算法实现难度和代码复杂性；具体实现由 Sidecar 代理完成，实现了逻辑与算法的分离，增加后续修改实现逻辑的灵活性；避免训练代码直接调用平台服务接口，避免了平台服务接口的外露。

以上已结合图 4 至图 8 描述了基于容器化应用管理平台进行 AI 模型的持续训练测评过程中所涉及的各种可能的操作和步骤。为了更清楚地说明该持续训练测评过程的整体流程，下面将参考图 9 进一步描述该持续训练测评过程。

图 9 示出了根据本申请的一些实施例的 AI 模型的持续训练测评任务流程的详细流程图。该持续训练测评任务流程可以在容器化应用管理平台上实现。这里，需要注意的是，在之前的描述中，主要是以图像处理领域的图像检测模型为示例来描述本申请所提供的训练测评任务流程（包括单次训练测评任务流程和持续测评任务流程），但是可以理解的是，该训练测评任务流程可以应用于各个领域的不同 AI 模型的训练和测评。对于不同领域的不同 AI 模型，平台可以根据相关领域的算法专家输入的与 AI 模型的训练测评任务相关联的流程配置信息，创建应用于容器化应用管理平台的流程任务模板。该流程任务模板可以包括：与 AI 模型的训练测评任务相关联的流程配置信息、以及训练测评任务中的各个步骤所对应的阶段标志。因而，在平台上可能存在与各个领域的不同 AI 模型相关联的多个流程任务模板，相应地，当要启动某个训练或测评任务时，需要查询该训练或测评任务所属的流程任务模板。

如图 9 所示，响应于接收到持续训练测评任务的启动命令，平台可以执行以下操作：查询流程任务模板，根据流程任务模板中的流程配置信息和阶段标志，对训练测评任务进行阶段拆分以构建训练阶段任务，根据训练阶段任务进行产物逃逸分析，以生成与训练阶段任务相关联的产物逃逸分析结果，基于与训练阶段任务相关联的产物逃逸结果，启动训练阶段任务。

在执行训练阶段任务期间，可以进行以下操作：进行模型训练迭代；在迭代过程中根据需要通过 SDK 通知 sidecar 代理对数据集预处理产物进行更新；当训练阶段任务达到预定的训练阶段性节点（例如，达到阶段性模型训练目标或预定义的训练迭代次数）时，在不中断所述训练阶段任务的情况下，通过 SDK 通知 sidecar 代理来启动测评阶段任务；利用 sidecar 代理根据阶段运行配置、产物逃逸分析结果，对逃逸产物进行存储；

sidecar 代理调用平台服务接口来启动测评阶段任务；继续进行后续模型训练。

如图 9 所示，平台可以在不中断训练阶段任务的情况下，在每次训练阶段任务达到预定的训练阶段性节点时启动相应的一个测评阶段任务，从而可以在平台上实现多个测评阶段任务（测评任务 1 至测评任务 n）的并发运行。而且，针对每个测评阶段任务，可以进行以下操作：查询流程任务模板；根据流程任务模板中的流程配置信息和阶段标志，构建测评阶段任务，根据测评阶段任务进行产物逃逸分析，以生成与测评阶段任务相关联的产物逃逸分析结果，并且基于与测评阶段任务相关联的产物逃逸结果，启动测评阶段任务。

综上，利用根据本申请的持续训练测评流程，可以实现不同领域的各种 AI 模型的训练测评功能的自动化和并行化。此外，采用 SDK 结合 Sidecar 代理的方式来实现产物文件的存储和传递以及平台服务接口的调用，使得训练任务容器不需要直接调用平台服务接口，避免了平台服务接口的外露。

图 10 示出了根据本申请实施例的 AI 模型的训练测评装置的硬件结构示意图。

该装置可以包括处理器 1001 以及存储有程序指令的存储器 1002。

具体地，上述处理器 1001 可以包括中央处理器（CPU），或者特定集成电路（Application Specific Integrated Circuit，ASIC），或者可以被配置成实施本申请实施例的一个或多个集成电路。

存储器 1002 可以包括用于数据或指令的大容量存储器。举例来说而非限制，存储器 1002 可包括硬盘驱动器（Hard Disk Drive，HDD）、软盘驱动器、闪存、光盘、磁光盘、磁带或通用串行总线（Universal Serial Bus，USB）驱动器或者两个或更多个以上这些的组合。在合适的情况下，存储器 1002 可包括可移除或不可移除(或固定)的介质。在合适的情况下，存储器 1002 可在综合网关容灾设备的内部或外部。在特定实施例中，存储器 1002 是非易失性固态存储器。

在特定实施例中，存储器 1002 包括只读存储器（ROM）。在合适的

情况下，该 ROM 可以是掩模编程的 ROM、可编程 ROM (PROM)、可擦除 PROM (EPROM)、电可擦除 PROM (EEPROM)、电可改写 ROM (EAROM) 或闪存或者两个或更多个以上这些的组合。

存储器可包括只读存储器 (ROM)，随机存取存储器 (RAM)，磁盘存储介质设备，光存储介质设备，闪存设备，电气、光学或其他物理/有形的存储设备。因此，通常，存储器包括一个或多个编码有包括计算机可执行指令的软件的有形 (非暂态) 可读存储介质 (例如，存储器设备)，并且当该软件被执行 (例如，由一个或多个处理器) 时，其可操作来执行参考根据本申请实施例的 AI 模型的训练测评方法所描述的操作。

处理器 1001 通过读取并执行存储器 1002 中存储的程序指令，以实现上述实施例中所述的 AI 模型的训练测评方法。

在一个示例中，AI 模型的训练测评装置还可包括通信接口 1003 和总线 1010。其中，如图 10 所示，处理器 1001、存储器 1002、通信接口 1003 通过总线 1010 连接并进行相互间的通信。

通信接口 1003，主要用于实现本申请实施例中各模块、装置、单元和/或设备之间的通信。

总线 1010 包括硬件、软件或两者，将装置的部件彼此耦接在一起。举例来说而非限制，总线可包括加速图形端口 (AGP) 或其他图形总线、增强工业标准架构 (EISA) 总线、前端总线 (FSB)、超传输 (HT) 互连、工业标准架构 (ISA) 总线、无限带宽互连、低引脚数 (LPC) 总线、存储器总线、微信道架构 (MCA) 总线、外围组件互连 (PCI) 总线、PCI-Express (PCI-X) 总线、串行高级技术附件 (SATA) 总线、视频电子标准协会局部 (VLB) 总线或其他合适的总线或者两个或更多个以上这些的组合。在合适的情况下，总线 1010 可包括一个或多个总线。尽管本申请实施例描述和示出了特定的总线，但本申请考虑任何合适的总线或互连。

结合上述实施例中的 AI 模型的训练测评方法，本申请实施例可由一种机器可读介质来实现。该机器可读介质上存储有程序指令；该程序指令被处理器执行时使得处理器执行上述实施例中所述的 AI 模型的训练测评

方法。

需要明确的是，本申请并不局限于上文所描述并在图中示出的特定配置和处理。为了简明起见，这里省略了对已知方法的详细描述。在上述实施例中，描述和示出了若干具体的步骤作为示例。但是，本申请的方法过程并不限于所描述和示出的具体步骤，本领域的技术人员可以在领会本申请的精神后，作出各种改变、修改和添加，或者改变步骤之间的顺序。

以上所述的结构框图中所示的功能块可以实现为硬件、软件、固件或者它们的组合。当以硬件方式实现时，其可以例如是电子电路、专用集成电路（ASIC）、适当的固件、插件、功能卡等等。当以软件方式实现时，本申请的元素是被用于执行所需任务的程序或者代码段。程序或者代码段可以存储在机器可读介质中，或者通过载波中携带的数据信号在传输介质或者通信链路上传送。“机器可读介质”可以包括能够存储或传输信息的任何介质。机器可读介质的例子包括电子电路、半导体存储器设备、ROM、闪存、可擦除 ROM（EROM）、软盘、CD-ROM、光盘、硬盘、光纤介质、射频（RF）链路，等等。代码段可以经由诸如因特网、内联网等的计算机网络被下载。

还需要说明的是，本申请中提及的示例性实施例，基于一系列的步骤描述了一些方法。但是，本申请不局限于上述步骤的顺序，也就是说，可以按照实施例中提及的顺序执行步骤，也可以不同于实施例中的顺序执行步骤，或者若干步骤同时执行。

以上参考根据本申请的实施例的方法和装置的流程图和/或框图，描述了本申请的各方面。应当理解，流程图和/或框图中的每个方框以及流程图和/或框图中各方框的组合可以由程序指令实现。这些程序指令可被提供给通用计算机、专用计算机、或其它可编程数据处理装置的处理器，以产生一种机器，使得经由计算机或其它可编程数据处理装置的处理器执行的这些指令使能对流程图和/或框图的一个或多个方框中指定的功能/动作的实现。这种处理器可以是但不限于是通用处理器、专用处理器、特殊应用处理器或者现场可编程逻辑电路。还可理解，框图和/或流程图中的每个方框以及框图和/或流程图中的方框的组合，也可以由执行指定的功能或动作的

专用硬件来实现，或可由专用硬件和计算机指令的组合来实现。

以上所述仅为本申请的具体实施方式，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，以上描述的装置的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。应理解，本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到各种等效的修改或替换，这些修改或替换都应涵盖在本申请的保护范围之内。

权 利 要 求 书

1. 一种用于人工智能 AI 模型的训练测评的装置，包括：
存储器，用于存储程序指令；以及
处理器，该处理器与所述存储器耦合，并且被配置为执行所述程序指令来：

创建应用于容器化应用管理平台的流程任务模板，该流程任务模板包括：与所述 AI 模型的训练测评任务相关联的流程配置信息、以及所述训练测评任务中的各个步骤所对应的阶段标志；并且

响应于接收到持续训练测评任务的启动命令，执行以下操作：

查询所述流程任务模板，

根据所述流程任务模板中的所述流程配置信息和所述阶段标志，
对所述训练测评任务进行阶段拆分以构建训练阶段任务，

根据所述训练阶段任务进行产物逃逸分析，以生成与所述训练阶段任务相关联的产物逃逸分析结果，

基于与所述训练阶段任务相关联的产物逃逸结果，启动所述训练阶段任务，并且

当所述训练阶段任务达到预定的训练阶段性节点时，在不中断所述训练阶段任务的情况下，启动一个或多个测评阶段任务的并发运行。

2. 根据权利要求 1 所述的装置，其中，启动一个或多个测评阶段任务的并发运行包括针对所述一个或多个测评阶段任务中的每个测评阶段任务来执行以下操作：

查询所述流程任务模板；

根据所述流程任务模板中的所述流程配置信息和所述阶段标志，构建所述测评阶段任务，

根据所述测评阶段任务进行产物逃逸分析，以生成与所述测评阶段任务相关联的产物逃逸分析结果，并且

基于与所述测评阶段任务相关联的产物逃逸结果，启动所述测评阶段

任务。

3. 根据权利要求 1 所述的装置，其中，所述预定的训练阶段性节点包括：所述训练阶段任务达到阶段性模型训练目标的节点、或者所述训练阶段任务达到预定义的训练迭代次数的节点。

4. 根据权利要求 1 所述的装置，其中，根据所述训练阶段任务进行产物逃逸分析包括：

基于所述流程任务模板，分析所述训练阶段任务中的各个步骤的产物在各个步骤之间的传递关系，所述产物包括输入产物和输出产物；

筛选出所述传递关系跨越所述训练测评任务中的不同阶段的产物，作为逃逸产物进行处理；并且

基于所述训练阶段任务中的各个步骤对应的阶段标志，建立所述逃逸产物的传递处理逻辑，作为与所述训练阶段任务相关联的产物逃逸分析结果。

5. 根据权利要求 2 所述的装置，其中，根据所述测评阶段任务进行产物逃逸分析包括：

基于所述流程任务模板，分析所述测评阶段任务中的各个步骤的产物在各个步骤之间的传递关系，所述产物包括输入产物和输出产物；

筛选出所述传递关系跨越所述训练测评任务中的不同阶段的产物，作为逃逸产物进行处理；并且

基于所述测评阶段任务中的各个步骤对应的阶段标志，建立所述逃逸产物的传递处理逻辑，作为与所述测评阶段任务相关联的产物逃逸分析结果。

6. 根据权利要求 1 至 5 中任一项所述的装置，其中，所述训练阶段任务被配置有关联的代理容器，并且启动所述训练测评任务中的测评阶段任务包括：利用所述代理容器，通过调用所述容器化应用管理平台的服务接口来启动所述一个或多个测评阶段任务的并发运行。

7. 根据权利要求 6 所述的装置，其中，所述操作还包括：利用所述代理容器根据所述产物逃逸分析结果存储产物文件，并辅助实现所述产物文件在所述训练阶段任务与所述一个或多个测评阶段任务之间的传递。

8. 根据权利要求 6 所述的装置，其中，所述训练阶段任务通过专用的软件开发工具包 SDK 与所述代理容器交互，以通知所述代理容器对与所述训练阶段任务相关联的产物进行更新，以及启动所述一个或多个测评阶段任务的并发运行。

9. 根据权利要求 6 所述的装置，其中，所述代理容器是基于用于所述容器化应用管理平台的 Sidecar 架构模式的 Sidecar 代理容器。

10. 根据权利要求 1 所述的装置，其中，所述处理器还被配置为执行所述程序指令来：

响应于接收到单次训练测评任务的启动命令，基于所述流程任务模板顺序执行所述 AI 模型的单次训练阶段任务和单次测评阶段任务。

11. 根据权利要求 1 所述的装置，其中，创建应用于容器化应用管理平台的流程任务模板包括：

基于与所述 AI 模型相关联的领域知识，构建所述 AI 模型的训练测评任务的任务步骤组成，该任务步骤组成包括完成所述 AI 模型的训练测评任务所需要的各个步骤；并且

为所述任务步骤组成中的各个步骤分配该步骤所属的阶段，以得到与该步骤相对应的阶段标志。

12. 根据权利要求 11 所述的装置，其中，所述 AI 模型包括图像检测模型，所述任务步骤组成包括图像预处理步骤、模型训练步骤、模型测评步骤和后处理步骤。

13. 一种用于人工智能 AI 模型的训练测评的方法，包括：

创建应用于容器化应用管理平台的流程任务模板，该流程任务模板包括：与所述 AI 模型的训练测评任务相关联的流程配置信息、以及所述训练测评任务中的各个步骤所对应的阶段标志；并且

响应于接收到持续训练测评任务的启动命令：

查询所述流程任务模板，

根据所述流程任务模板中的所述流程配置信息和所述阶段标志，
对所述训练测评任务进行阶段拆分以构建训练阶段任务，

根据所述训练阶段任务进行产物逃逸分析，以生成与所述训练阶

段任务相关联的产物逃逸分析结果，

基于与所述训练阶段任务相关联的产物逃逸结果，启动所述训练阶段任务，并且

当所述训练阶段任务达到预定的训练阶段性节点时，在不中断所述训练阶段任务的情况下，启动一个或多个测评阶段任务的并发运行。

14. 根据权利要求 13 所述的方法，其中，启动一个或多个测评阶段任务的并发运行包括针对所述一个或多个测评阶段任务中的每个测评阶段任务来执行以下操作：

查询所述流程任务模板；

根据所述流程任务模板中的所述流程配置信息和所述阶段标志，构建所述测评阶段任务，

根据所述测评阶段任务进行产物逃逸分析，以生成与所述测评阶段任务相关联的产物逃逸分析结果，并且

基于与所述测评阶段任务相关联的产物逃逸结果，启动所述测评阶段任务。

15. 根据权利要求 13 所述的方法，其中，所述预定的训练阶段性节点包括：所述训练阶段任务达到阶段性模型训练目标的节点、或者所述训练阶段任务达到预定义的训练迭代次数的节点。

16. 根据权利要求 13 所述的方法，其中，根据所述训练阶段任务进行产物逃逸分析包括：

基于所述流程任务模板，分析所述训练阶段任务中的各个步骤的产物在各个步骤之间的传递关系，所述产物包括输入产物和输出产物；

筛选出所述传递关系跨越所述训练测评任务中的不同阶段的产物，作为逃逸产物进行处理；并且

基于所述训练阶段任务中的各个步骤对应的阶段标志，建立所述逃逸产物的传递处理逻辑，作为与所述训练阶段任务相关联的产物逃逸分析结果。

17. 根据权利要求 14 所述的方法，其中，根据所述测评阶段任务进行

产物逃逸分析包括：

基于所述流程任务模板，分析所述测评阶段任务中的各个步骤的产物在各个步骤之间的传递关系，所述产物包括输入产物和输出产物；

筛选出所述传递关系跨越所述训练测评任务中的不同阶段的产物，作为逃逸产物进行处理；并且

基于所述测评阶段任务中的各个步骤对应的阶段标志，建立所述逃逸产物的传递处理逻辑，作为与所述测评阶段任务相关联的产物逃逸分析结果。

18. 根据权利要求 13 至 17 中任一项所述的方法，其中，所述训练阶段任务被配置有关联的代理容器，并且启动所述训练测评任务中的测评阶段任务包括：利用所述代理容器，通过调用所述容器化应用管理平台的服务接口来启动所述一个或多个测评阶段任务的并发运行。

19. 根据权利要求 18 所述的方法，还包括：利用所述代理容器根据所述产物逃逸分析结果存储产物文件，并辅助实现所述产物文件在所述训练阶段任务与所述一个或多个测评阶段任务之间的传递。

20. 根据权利要求 18 所述的方法，其中，所述训练阶段任务通过专用的软件开发工具包 SDK 与所述代理容器交互，以通知所述代理容器对与所述训练阶段任务相关联的产物进行更新，以及启动所述一个或多个测评阶段任务的并发运行。

21. 根据权利要求 18 所述的方法，其中，所述代理容器是基于用于所述容器化应用管理平台的 Sidecar 架构模式的 Sidecar 代理容器。

22. 根据权利要求 13 所述的方法，还包括：

响应于接收到单次训练测评任务的启动命令，基于所述流程任务模板顺序执行所述 AI 模型的单次训练阶段任务和单次测评阶段任务。

23. 根据权利要求 13 所述的方法，其中，创建应用于容器化应用管理平台的流程任务模板包括：

基于与所述 AI 模型相关联的领域知识，构建所述 AI 模型的训练测评任务的任务步骤组成，该任务步骤组成包括完成所述 AI 模型的训练测评任务所需要的各个步骤；并且

为所述任务步骤组成中的各个步骤分配该步骤所属的阶段，以得到与该步骤相对应的阶段标志。

24. 根据权利要求 23 所述的方法，其中，所述 AI 模型包括图像检测模型，所述任务步骤组成包括图像预处理步骤、模型训练步骤、模型测评步骤和后处理步骤。

25. 一种机器可读介质，其上存储有程序指令，所述程序指令在被处理器执行时，使得所述处理器执行如权利要求 13 至 24 中任一项所述的用于人工智能 AI 模型的训练测评的方法。

26. 一种用于人工智能 AI 模型的训练测评的设备，包括用于执行如权利要求 13 至 24 中任一项所述的用于 AI 模型的训练测评的方法的装置。