

用于检验申报信息真实性的方法、系统、设备及介质

技术领域

本公开涉及人工智能技术领域，更具体地，涉及一种用于检验申报信息真实性的方法、装置、设备及介质。

背景技术

在海关、空运、物流等物资运输或交割环节中，经常需要查验申报人所申报的物资与实际运送的物资是否真实吻合，以杜绝走私之类事件的发生。然而，在运输途中物品通常装在集装箱等密封的容器内，不易被观测，会给查验环节带来极大不便。

现有查验方法主要有两种：一种是对物品逐件开箱、对物品的名称、规格型号、数量、重量、产地信息进行逐一核实，然而这样会消耗大量人力；另一种是按照一定比例对物品进行抽查、核验报关信息，但这样难以做到“一一开箱”，可能漏掉对关键信息的查验。为了提高通关效率，现有技术中也有引入 X 射线探测装置进行查验的方式。

在实现本公开构思的过程中，发明人发现现有技术中至少存在如下问题：通过 X 射线探测装置扫描物品进行查验时，由于同一申报类别（例如，申报的税号类别，在海关报关单中为商品编号）下的物品可能包括很多不同名称、不同规格型号的货物，这些物品在 X 射线图像（以下均称为机检辐射图像）上的表现差异较大，增加了查验难度。另外，在海关等业务中经常出现拼箱、拼单或是多种不同品类或名称的物品混装，会导致申报单和集装箱并不是一对一的简单映射关系，这样单独通过机检辐射图像进行查验，会存在大量名称、规格型号和图像无法对齐的复杂情形，给查验工作的准确性和便捷性带来很大困难。

发明内容

有鉴于此，本公开实施例提供了一种联合机检辐射图像和申报信息进行智能查验的用于检验申报信息真实性的方法、装置、设备及介质，可以提高查验准确率，提升查验效率。

本公开实施例的一个方面，提供了一种用于检验申报信息真实性的方法。所述方法包括：获取对装载了物品的容器进行扫描而获得的机检辐射图像；获取对所述容器内的物品进行申报的申报信息；对所述机检辐射图像中的物品的图像信息进行识别，得到与所述机检辐射图像对应的图像特征；对所述申报信息中的物品的文本信息进行识别，得到与所述申报信息对应的文本特征，其中，所述文本特征用于表征所述申报信息中的物品所属的申报类别；以所述图像特征作为输入信息，以所述文本特征作为外部引入特征，对所述容器中的物品的申报类别进行甄别；以及当所述容器中至少一个物品的申报类别不属于所述申报信息中的申报类别时，确定所述申报信息存疑。

根据本公开的实施例，所述图像特征包括与所述机检辐射图像中的N1个物品的图像信息分别对应的N1个第一图像特征向量，其中，N1为大于或等于1的整数。

根据本公开的实施例，所述对所述机检辐射图像中的物品的图像信息进行识别，得到与所述机检辐射图像对应的图像特征包括：利用目标检测算法将所述机检辐射图像中的不同物品划分成独立的图像块，得到N1个图像块；提取每个所述图像块对应的第二图像特征向量；以及基于每个所述图像块对应的所述第二图像特征向量，得到与所述图像块所表示的物品的图像信息对应的所述第一图像特征向量。

根据本公开的实施例，所述提取每个所述图像块对应的第二图像特征向量包括：利用图像特征提取模块对每个所述图像块进行图像识别，得到每个所述图像块对应的所述第二图像特征向量；其中，所述图像特征提取模块包括卷积神经网络。

根据本公开的实施例，所述图像特征提取模块包括以resnet作为基础网络，在resnet池化层后添加SE-block的网络结构。

根据本公开的实施例，所述基于每个所述图像块对应的所述第二图像特征向量，得到与所述图像块所表示的物品的图像信息对应的所述第一图

像特征向量包括：获取每个所述图像块在所述机检辐射图像中的位置信息；以及基于同一个所述图像块对应的所述第二图像特征向量和所述位置信息，得到所述第一图像特征向量。

根据本公开的实施例，所述基于同一个所述图像块对应的所述第二图像特征向量和所述位置信息，得到所述第一图像特征向量包括：利用所述图像块的所述位置信息处理所述第二图像特征向量；将处理后的所述第二图像特征向量输入编码器；以及获得所述编码器的输出，以得到与所述图像块对应的所述第一图像特征向量。

根据本公开的实施例，所述以所述图像特征作为输入信息，以所述文本特征作为外部引入特征，对所述容器中的物品的申报类别进行甄别包括：以所述图像特征作为跨模态解码器的输入信息，以所述文本特征作为所述跨模态解码器的注意力机制的外部引入特征，利用所述跨模态解码器对所述容器中的物品的申报类别进行甄别。

根据本公开的实施例，所述编码器与所述跨模态解码器联合训练。

根据本公开的实施例，所述编码器采用 **transformer** 编码器模型。

根据本公开的实施例，所述跨模态解码器采用 **transformer** 解码器模型。

根据本公开的实施例，所述文本特征包括与所述申报信息中的 N_2 种物品分别对应的文本特征向量，其中 N_2 为大于或等于 1 的整数。

根据本公开的实施例，所述对所述申报信息中的物品的文本信息进行识别，得到与所述申报信息对应的文本特征包括：提取所述申报信息中每一种物品的名称信息和规格型号信息；对于每一种物品，将所述名称信息处理为第一语句，将所述规格型号信息处理为第二语句；将对应于同一种物品的所述第一语句和所述第二语句作为文本特征提取投影模块的输入，利用所述文本特征提取投影模块对该物品所属的申报类别进行分类；以及以所述文本特征提取投影模块针对每一种物品的类别输出结果作为该物品对应所述文本特征向量；其中，针对 N_2 种物品对应得到 N_2 个所述文本特征向量。

根据本公开的实施例，所述文本特征提取投影模块采用 **BERT** 模型。

本公开实施例的第二方面，提供了一种用于检验申报信息真实性的系统。所述系统包括信息获取子系统、特征提取子系统、特征融合子系统以

及结论判定子系统。信息获取子系统用于获取对装载了物品的容器进行扫描而获得的机检辐射图像，以及获取对所述容器内的物品进行申报的申报信息。特征提取子系统用于对所述机检辐射图像中的物品的图像信息进行识别，得到与所述机检辐射图像对应的图像特征；以及对所述申报信息中的物品的文本信息进行识别，得到与所述申报信息对应的文本特征，其中，所述文本特征用于表征所述申报信息中的物品所属的申报类别。特征融合子系统用于以所述图像特征作为输入信息，以所述文本特征作为外部引入特征，对所述容器中的物品的申报类别进行甄别。结论判定子系统用于当所述容器中至少一个物品的申报类别不属于所述申报信息中的申报类别时，确定所述申报信息存疑。

根据本公开的实施例，所述图像特征包括与所述机检辐射图像中的 $N1$ 个物品的图像信息分别对应的 $N1$ 个第一图像特征向量，其中， $N1$ 为大于或等于 1 的整数。根据本公开的实施例，所述特征提取子系统包括图像预处理模块、图像特征提取模块以及图像特征映射模块。图像预处理模块用于利用目标检测算法将所述机检辐射图像中的不同物品划分成独立的图像块，得到 $N1$ 个图像块。图像特征提取模块用于提取每个所述图像块对应的第二图像特征向量。图像特征映射模块用于基于每个所述图像块对应的所述第二图像特征向量，得到与所述图像块所表示的物品的图像信息对应的所述第一图像特征向量。

根据本公开的实施例，所述文本特征包括与所述申报信息中的 $N2$ 种物品分别对应的文本特征向量，其中 $N2$ 为大于或等于 1 的整数。根据本公开的实施例，所述特征提取子系统包括申报信息预处理模块和文本特征提取投影模块。所述申报信息预处理模块用于提取所述申报信息中每一种物品的名称信息和规格型号信息，对于每一种物品，将所述名称信息处理为第一语句，将所述规格型号信息处理为第二语句。所述文本特征提取投影模块用于以对应于同一种物品的所述第一语句和所述第二语句作为输入，对该物品所属的申报类别进行分类；以及以所述文本特征提取投影模块针对每一种物品的类别输出结果作为该物品对应所述文本特征向量；其中，针对 $N2$ 种物品对应得到 $N2$ 个所述文本特征向量。

本公开实施例的另一方面，提供了一种电子设备。所述电子设备包括一个或多个存储器、以及一个或多个处理器。所述存储器存储有可执行指令。所述处理器执行所述可执行指令以实现如上所述的方法。

本公开实施例的另一方面，提供了一种计算机可读存储介质，存储有计算机可执行指令，所述指令在被执行时用于实现如上所述的方法。

本公开实施例的另一方面，提供了一种计算机程序，所述计算机程序包括计算机可执行指令，所述指令在被执行时用于实现如上所述的方法。

上述一个或多个实施例具有如下优点或益效果：通过联合文本形式的申报信息以及机检辐射图像，利用基于注意力机制的跨模态融合器将图像特征与申报信息文本特征融合，实现对通关物品的智能查验，提高对申报信息的查验准确率。

附图说明

通过以下参照附图对本公开实施例的描述，本公开的上述以及其他目的、特征和优点将更为清楚，在附图中：

图1示意性示出了根据本公开实施例的用于检验申报信息真实性的方法、系统、设备、介质和程序产品的应用场景；

图2示意性示出了根据本公开实施例的用于检验申报信息真实性的方法的构思示意图；

图3示意性示出了根据本公开实施例的用于检验申报信息真实性的方法的流程图；

图4示意性示出了根据本公开实施例的用于检验申报信息真实性的方法中提取图像特征的流程图；

图5示意性根据本公开实施例的用于检验申报信息真实性的方法中结合位置信息得到图像特征的流程图；

图6示意性示出了根据本公开实施例的用于检验申报信息真实性的方法中提取图像特征的流程示意；

图7示意性示出了根据本公开实施例的用于检验申报信息真实性的方法中提取文本特征的流程图；

图 8 示意性示出了根据本公开实施例的用于检验申报信息真实性的方法中提取文本特征的流程示意；

图 9 示意性示出了根据本公开实施例的用于检验申报信息真实性的方法中对所述容器中的物品的申报类别进行甄别的流程示意；

图 10 示意性示出了根据本公开实施例的用于检验申报信息真实性的系统的框图；

图 11 示意性示出了根据本公开实施例的用于检验申报信息真实性的系统中特征提取子模块的框图；

图 12 示意性示出了根据本公开另一实施例的用于检验申报信息真实性的系统的整体结构示意；

图 13 示意性示出了根据本公开一实施例的用于检验申报信息真实性的系统中特征提取子系统的结构示意；

图 14 示意性示出了根据本公开一实施例的用于检验申报信息真实性的系统中跨模态解码器的结构示意；以及

图 15 示意性示出了适于实现根据本公开实施例的用于检验申报信息真实性的方法的电子设备的框图。

具体实施方式

以下，将参照附图来描述本公开的实施例。但是应该理解，这些描述只是示例性的，而并非要限制本公开的范围。在下面的详细描述中，为便于解释，阐述了许多具体的细节以提供对本公开实施例的全面理解。然而，明显地，一个或多个实施例在没有这些具体细节的情况下也可以被实施。此外，在以下说明中，省略了对公知结构和技术的描述，以避免不必要地混淆本公开的概念。

在本文中，需要理解的是，说明书及附图中的任何元素数量均用于示例而非限制，以及任何命名（例如，第一、第二）都仅用于区分，而不具有任何限制含义。

为了能够有效地核实封装在容器内的物品与申报信息是否一致，本公开的实施例提供了一种用于检验申报信息真实性的方法、系统、设备、介质和程序产品，联合文本形式的申报信息以及机检辐射图像，利用基于注

注意力机制的跨模态融合器将图像特征与申报信息文本特征融合，实现对通关物品的智能查验，确定是否存在伪瞒报行为。

具体地，本公开的实施例提供了一种用于检验申报信息真实性的方法、系统、设备、介质和方法。该方法包括首先获取对装载了物品的容器进行扫描而获得的机检辐射图像，以及获取对容器内的物品进行申报的申报信息；然后对机检辐射图像中的物品的图像信息进行识别，得到与机检辐射图像对应的图像特征，以及对申报信息中的物品的文本信息进行识别，得到与申报信息对应的文本特征，其中，文本特征可以映射申报信息中的物品所属的申报类别；接下来以图像特征作为输入信息，以文本特征作为外部引入特征，对所述容器中的物品的申报类别进行甄别，并且当甄别出容器中至少一个物品的申报类别不属于申报信息中的申报类别时，确定申报信息存疑。

图1示意性示出了根据本公开实施例的用于检验申报信息真实性的方法、系统、设备、介质和程序产品的应用场景100。需要注意的是，图1所示仅为可以应用本公开实施例的系统架构的示例，以帮助本领域技术人员理解本公开的技术内容，但并不意味着本公开实施例不可以用于其他设备、系统、环境或场景。

如图1所示，根据该实施例的应用场景100可以包括X射线探测装置101、服务器102和终端设备103。其中，X射线探测装置101、终端设备103分别与服务器102通信连接。

容器12中装载物品在通过X射线探测装置101时，可以被X射线探测装置101扫描得到机检辐射图像。

X射线探测装置101可以扫描得到的机检辐射图像上传到服务器102。与此同时，服务器102也可以通过从数据库搜索、或云端搜索、或者工作人员11通过终端设备103上传等方式，获取到与容器12中装载物品对应的申报信息（例如，一个或多个申报单）。根据本公开的实施例，服务器102可以执行根据本公开实施例的用于检验申报信息真实性的方法，联合机检辐射图像和申报信息，利用基于注意力机制的跨模态融合器，来分析容器12中的物品与申报信息是否一致，是否存在瞒报的嫌疑。然后服务器102可以将分析结果发送给终端设备103，展示给工作人员11。然后工

作人员 11 可以根据分析结果,若对容器 12 中的物品的申报信息存有嫌疑,则可以对容器 12 进行开箱查验;若对容器 12 中的物品的申报信息没有嫌疑,则可以选择对容器 12 抽检或放行。这样在节约查验环节的人力投入的同时,还能够在一定程度上提高查验的全面性和准确性。

在实际应用中,由于可能存在拼单或拼箱的现象,容器 12 内的物品可以对应一个申报单,也可以对应多个申报单,即容器 12 内的物品可以是多个申报单中的物品拼装在一起形成的。在另一些情形下,也可以是 x 个容器中装载有 y 个申报单申报物品,其中, x 、 y 均为整数,且 $x \neq y$ 。在这种情形下,可以将 x 个容器中的物品连续扫描得到机检辐射图像,与 y 个申报单中的申报信息相对应进行处理。或者,在这种情形下,也可以将对 x 个容器其中每个容器中的物品扫描得到的机检辐射图像,与该 y 个报关单中的申报信息对应进行处理。

需要说明的是,本公开实施例所提供的用于检验申报信息真实性的方法一般可以由服务器 102 执行。相应地,本公开实施例所提供的用于检验申报信息真实性的系统、设备、介质和程序产品一般可以设置于服务器 102 中。本公开实施例所提供的用于检验申报信息真实性的方法也可以由不同于服务器 102 且能够与 X 射线探测装置 101、终端设备 103 和/或服务器 102 通信的服务器或服务器集群执行。相应地,本公开实施例所提供的用于检验申报信息真实性的系统、设备、介质和程序产品也可以设置于不同于服务器 102 且能够与 X 射线探测装置 101、终端设备 103 和/或服务器 102 通信的服务器或服务器集群中。

应该理解,图 1 中的设备的数目和种类仅仅是示意性的。根据实现需要,可以具有任意数目和种类的设备。

图 2 示意性示出了根据本公开实施例的用于检验申报信息真实性的方法的构思示意图。

如图 2 所示,根据本公开的实施例,为了能够有效地核实容器 20 内装载的物品与申报信息 21 中申报的物品类别是否一致,可以从申报信息 21 中获取申报的物品的名称信息以及规格型号等描述性信息,联合对容器 20 扫描得到的机检辐射图像 22,利用根据本公开实施例的用于检验申报

信息真实性的系统 201，通过自然语言处理技术、深度学习等方法，实现对容器 20 内的物品的智能查验，确定是否存在瞒报行为。

根据本公开的一些实施例，可以将机检辐射图像以物品为单位，划分成不同的图像块，在系统 201 中利用图像识别等视觉编码技术获取图像特征，与此同时将从申报信息中提取的文本特征作为注意力机制的外部引入特征，利用跨模态分类器将图像特征与文本融合，通过多类别约束得到每个图像块的分类结果，并将该结果与申报信息中填写的申报类别进行比较，得到最终查验结果。

其中，该用于检验申报信息真实性的系统 201 可以被实现为例如下文所描述的系统 1000、系统 1200、或电子设备 1500，计算机存储介质，或者程序产品的其中之一或任意的结合，用于实现根据本公开实施例的用于检验申报信息真实性的方法。

图 3 示意性示出了根据本公开实施例的用于检验申报信息真实性的方法的流程图。

如图 3 所示，根据该实施例的用于检验申报信息真实性的方法可以包括操作 S310~操作 S360。

首先在操作 S310，获取对装载了物品的容器进行扫描而获得的机检辐射图像。

并且在操作 S320，获取对容器内的物品进行申报的申报信息。

接下来在操作 S330，对机检辐射图像中的物品的图像信息进行识别，得到与机检辐射图像对应的图像特征。

并且在操作 S340，对申报信息中的物品的文本信息进行识别，得到与申报信息对应的文本特征。其中，文本特征可以映射到申报信息中的物品所属的申报类别。例如，根据申报信息中申报的物品的描述性信息（例如，名称信息以及规格型号等），对所申报的物品所属的税号类别或者商品类别进行分类。

然后在操作 S350，以所述图像特征作为输入信息，以所述文本特征作为外部引入特征，对所述容器中的物品的申报类别进行甄别。

在一个实施例中，可以以图像特征作为跨模态解码器的输入信息，以文本特征作为跨模态解码器的注意力机制的外部引入特征，利用跨模态解码器对容器中的物品的申报类别进行甄别。

跨模态解码器例如可以采用 transformer 解码器时，可以以文本特征作为外部引入特征引入到 transformer 解码器的互注意力模块中，以图像特征作为 transformer 解码器的输入，对图像特征中各个物品所属的申报类别进行甄别。通过文本特征和图像特征之间的交互操作，完成跨模态的信息融合。

此后在操作 S360，当容器中至少一个物品的申报类别不属于申报信息中的申报类别时，确定申报信息存疑。此处申报信息中的申报类别可以是用户填写申报单时所填写的申报类别。用户填写申报单时所填写的申报类别与文本特征可以映射到申报类别可能相同也可能不同，毕竟用户填写申报时可能存在舞弊情况。因此，本公开实施例通过文本特征来映射物品真实的申报类别，并通过文本特征和图像特征的融合输出，对用户填写的申报信息进行核验。

图像特征可以包括与机检辐射图像中的 N1 个物品的图像信息分别对应的 N1 个第一图像特征向量，其中，N1 为大于或等于 1 的整数。例如，在提取图像特征时，在机检辐射图像中根据物品的轮廓边界等将机检辐射图像划分为不同的图像块，然后可以针对每一个图像块识别出第一图像特征向量。

文本特征也可以包括与申报信息中的 N2 种物品分别对应的 N2 个文本特征向量，其中 N2 为大于或等于 1 的整数。本公开实施例根据申报信息中的物品描述信息对物品所属的申报类别进行识别，从而，当申报有 N2 种物品时，会得到 N2 个文本特征向量。

根据本公开的实施例，当将 N1 个第一图像特征向量分别或者组合成序列输入到跨模态解码器中，并将 N2 第二特征向量组合成序列输入到跨模态解码器的互注意力模块中，通过特征之间的交互操作，可以实现跨模态的信息融合。其中，每个第一特征向量与 N2 个第二特征向量交互后，都会得到一个对应的结果，该结果为容器中的物品对应的每个类别的概率输出。即每个物品会得到其预测类别的概率值。可以采用 topN（例如，

N=1 或 2) 的方式将排名靠前的 N 个类别作为该物品的候选类别, 当候选类别均不包含在申报信息中所填写的申报类别中时, 认为风险较大, 申报信息存在嫌疑。否则, 可以认为申报信息无嫌疑。

由此可见, 本公开的实施例通过联合申报信息与机检辐射图像对申报物品进行综合分析, 提高了信息使用效率和查验准确率。

图 4 示意性示出了根据本公开实施例的用于检验申报信息真实性的方法中操作 S330 提取图像特征流程图。

如图 4 所示, 根据该实施操作 S330 可以包括操作 S331~操作 S333。

在操作 S331, 利用目标检测算法将机检辐射图像中的不同物品划分成独立的图像块, 得到 N1 个图像块。

在操作 S332, 提取每个图像块对应的第二图像特征向量。例如, 可以利用图像特征提取模块来提取图像块对应的第二图像特征向量。在一个实施例中, 图像特征提取模块可以是卷积神经网络。

根据本公开的一个实施例, 图像特征提取模块具体可以是以 resnet 作为基础网络, 在 resnet 池化层后添加 SE-block 的网络结构。例如, 可以采用 resnet 作为基础网络, 并在原有 resnet 池化层后添加 SE-block (Squeeze-Extract Block)、全连接层 fc1 和全连接层 fc2。在一个实施例中, 可以将全连接层 fc1 的输出结果作为第二图像特征向量, 由此获取一组第二图像特征向量。

图像特征提取模块可以利用历史上申报的机检辐射图像中的图像块进行训练。例如, 获取同一物品不同形态下的图像块, 通过预先标注, 让图像特征提取模块来学习和识别同一物品在机检辐射图像中可能的形态。

在操作 S333, 基于每个图像块对应的第二图像特征向量, 得到与图像块所表示的物品的图像信息对应的第一图像特征向量。

在一个实施例中, 可以直接将第二图像特征向量作为第一特征向量。

在另一些实施例中, 也可以对第二图像特征向量进一步处理, 得到第一图像特征向量。例如, 可以根据每个图像块在机检辐射图像中的位置信息, 对第二特征向量进行处理, 得到与能够反映图像块在机检辐射图像中的空间位置关系的特征, 作为第一特征向量。这样, 在操作 S350 中进行

类别识别时，可以更全面，也可以一定程度上避免对容器内的物品的漏识别或重复识别等，提高识别效率。

图 5 示意性根据本公开实施例的用于检验申报信息真实性的方法中在操作 S333 结合位置信息得到图像特征的流程图。

如图 5 所示，根据本公开实施例操作 S333 可以包括操作 S501 和操作 S502。在该实施例中可以根据每个图像块的位置信息对第二特征向量进行处理，得到作为第一特征向量。

具体地在操作 S501，获取每个图像块在机检辐射图像中的位置信息。该位置信息例如可以是每个图像块外轮廓中的定位点坐标。比如，当图像块为规则的几何图形时（通常海关申报等业务中物品都被封装在包装盒中），位置信息可以是几何图形的顶点或交点的坐标。

然后在操作 S502，基于同一个图像块对应的第二图像特征向量和位置信息，得到第一图像特征向量。

在一个实施例中，首先可以利用图像块的位置信息处理第二图像特征向量，例如可以将图像块的位置信息处理为向量，然后与图像块对应的第二图像特征向量连接起来；或者，例如可以利用图像块的位置信息对第二特征向量进行编码等映射转换处理。然后将利用位置信息处理后的第二图像特征向量输入到编码器，并获得编码器的输出，从而可以以编码器的输出作为到与图像块对应的第一图像特征向量。

根据本公开的实施例，在以编码器的输出作为与图像块对应的第一图像特征向量时，该编码器和操作 S350 中所使用的跨模态解码器可以联合训练。

联合训练时，编码器和跨模态解码器构成上下游关系，其中，可以以历史上申报的各类物品的图像块的位置信息对第二特征向量进行处理后的向量作为编码器的输入，将编码器的输出作为跨模态解码器的输入，同时将历史申报信息对应的文本特征作为跨模态解码器的互注意力模块的输入，获取跨模态解码器输出的图像块的申报类别，然后基于跨模态解码器输出的图像块的申报类别、与对该图像块标注的申报类别之间的误差，反复训练编码器和跨模态融合解码器。

根据本公开的实施例，编码器可以采用 transformer 编码器模型。相应地，跨模态解码器可以采用 transformer 解码器模型。

图 6 示意性示出了根据本公开实施例的用于检验申报信息真实性的方法中提取图像特征的流程示意。

结合图 2 和图 6 所示，可以利用目标检测法，根据物品在机检辐射图像 22 中的表现，划分成独立的图像块，并提取相应位置坐标。

然后利用图像特征提取模块 601 提取不同图像块的特征。在图 6 中图像特征提取模块 601 可以采用 resnet（深度残差网络，一种经典的卷积神经网络）作为基础网络，并在原有 resnet 池化层后添加 SE-block（Squeeze-Extract Block）神经网络结构单元、全连接层 fc1-relu（relu 为线性整流单元）和全连接层 fc2 层。其中，可以将 SE-block 通道数目设定为类别数目，以此方式引入通道注意力机制，并将全连接层 fc1 的输出结果作为每个图像块的第二图像特征向量，由此获取一组第二图像特征向量。假定其大小为 $N1 \times \text{dim}$ ，其中 N1 表示图像块的个数，dim 表示每个第二图像特征向量的特征维度。

接下来可以将 N1 个第二图像特征向量（即， $N1 \times \text{dim}$ ）与每个图像块对应的位置信息代入到 transformer 编码器 601 中，得到 N1 个第一图像特征向量，然后以该结果作为操作 S350 中跨模态解码器（例如，transformer 解码器）的输入。

图 7 示意性示出了根据本公开实施例的用于检验申报信息真实性的方法中在操作 S340 提取文本特征的流程示意图。

如图 7 所示，根据本公开的实施例操作 S340 可以包括操作 S341~操作 S344。

在操作 S341，提取申报信息中每一种物品的名称信息和规格型号信息。

在操作 S342，对于每一种物品，将名称信息处理为第一语句，将规格型号信息处理为第二语句。

在操作 S343，将对应于同一种物品的第一语句和第二语句作为文本特征提取投影模块的输入，利用文本特征提取投影模块对该物品所属的申报类别进行分类。

在操作 S344, 以文本特征提取投影模块针对每一种物品的类别输出结果作为该物品对应文本特征向量, 其中, 针对 $N2$ 种物品对应得到 $N2$ 个文本特征向量。

在本公开的一个实施例中, 文本特征提取投影模块可以采用 BERT 模型。在训练时, 可以搜集历史申报信息来训练 BERT 模型。将历史申报信息中每种物品的名称信息处理为第一语句, 将规格型号信息处理为第二语句, 形成语句序列输入到 BERT 模型, 输出每种物品的类别。然后, 基于 BERT 模型输出的文本特征向量所对应的类别, 与该每种物品实际所属的申报类别之间的误差, 对 BERT 模型 (即, Bidirectional Encoder Representations from Transformers 模型) 进行反复训练。

图 8 示意性示出了根据本公开实施例的用于检验申报信息真实性的方法中提取文本特征的流程示意。

结合图 2 和图 8, 该实施例中采用 BERT 模型, 提取申报信息 21 中的文本特征。

如果申报信息 21 中只有一种物品。则从该申报信息 21 中直接提取该物品的名称信息和规格型号, 预处理后利用 BERT 模型 801 提取将类别输出作为该商品文本特征, 即文本特征向量, 维度为 $1 \times \text{dim}$ 。

如果申报信息 21 包括多种物品, 假设 $N2$ 种。 $N2$ 表示物品名称的个数, dim 为特征维度。对于多种物品, 需要先将同一种物品的名称信息和规格型号逐一对齐, 再重复上述步骤, 得到 $N2$ 个 dim 维文本特征向量, 构成 $N2 \times \text{dim}$ 的文本特征向量序列。

以此方式, 利用自然语言处理方法获得能够表示物品名称信息和规格型号的文本特征。其中, 将文本特征作为外部引入特征引入到跨模块态解码器互注意力模块中。

图 9 示意性示出了根据本公开实施例的用于检验申报信息真实性的方法中在操作 S350 对容器中的物品的申报类别进行甄别的流程示意。

如图 9 所示, 操作 S350 中可以采用跨模态解码器 901 对容器中的物品的申报类别进行甄别。其中, 跨模态解码器可以采用 transformer 解码器 901。可以将 $N1$ 个第一图像特征向量作为 transformer 解码器 901 的输入, 将 $N2$ 个文本特征向量作为 transformer 解码器 901 的互注意力模块的

输入，取 transformer 解码器 901 的输出值，作为最终比对结果。即每个物品块会得到其预测的申报类别的概率值，采用 topN 的方式将排名靠前的 N 个类别作为该图像块的候选类别，当候选类别均不属于申报类别时，则认为风险较大，否则，可以认为无嫌疑。

为了便于计算，可以设置图像特征和文本特征中向量的维度 dim 相等。

对于在同一申报类别下具有多种品类物品的情形，单独利用图像信息很难进行准确判断。本公开实施例在跨模态解码器中，通过机检辐射图像与申报信息的交互，利用自然语言处理方法获得能够表示物品名称信息和规格型号的文本特征，可以从物品文本描述信息方面来弥补视图像信息的不足，提高识别准确率。

在使用 transformer 解码器 901 进行跨模态融合时，N2 个文本特征向量作为 transformer 解码器 901 的互注意力模块的外部特征，需要整体输入。N1 个第一图像特征向量为 transformer 解码器 901 的识别对象，可以分别输入，或者组合成序列输入，对此可以根据实际设置。

根据本公开的实施例，为了应对在一种申报类别对应多种物品的情形下的准确报关查验问题，采用目标检测法将原来的整张机检辐射图像分割成为不同的图像块，通过对图像块的特征提取得到独立的物品特征，可以提高识别准确率。

为了能够有效提取物品特征，本公开实施例在图像特征提取模块中可以将 SE-block 通道数目设置为类别数目以引入通道注意力机制，提高图像特征提取的准确率。

为了体现物品块在空间上的位置关系，本公开实施例可以将图像块的位置信息与特征信息，一并送入到 transformer 编码器中，得到具有空间位置关系的特征（即，第一特征向量）。

本公开实施例，在模型训练过程中，采用了三种训练任务相结合的方式，即通过图像训练来训练图像特征提取模块（如，模块 601）、通过文本训练来训练文本特征提取投影模块（如，BERT 模型 801），以及通过图像-文本训练来联合训练编码器和跨模态解码器（如，transformer 编码器 602 和 transformer 编码器 901），使得不同任务之间达到了相辅相成的效果。

图 10 示意性示出了根据本公开实施例的用于检验申报信息真实性的系统 1000 的框图。

如图 10 所示，根据本公开实施例的用于检验申报信息真实性的系统 1000 可以包括信息获取子系统 110、特征提取子系统 120、特征融合子系统 130 以及结论判定子系统 140。该系统 1000 可以用于实现参考图 3~图 9 所描述的方法。

具体地，信息获取子系统 110 可以用于获取对装载了物品的容器进行扫描而获得的机检辐射图像，以及获取对容器内的物品进行申报的申报信息。在一个实施例中，信息获取子系统 110 可以用于执行操作 S310 和操作 S320。

特征提取子系统 120 可以用于对机检辐射图像中的物品的图像信息进行识别，得到与机检辐射图像对应的图像特征，以及对申报信息中的物品的文本信息进行识别，得到与申报信息对应的文本特征，其中，文本特征用于表征申报信息中的物品所属的申报类别。在一个实施例中，特征提取子系统 120 可以用于执行操作 S330 和操作 S340。

特征融合子系统 130 可以用于以所述图像特征作为输入信息，以所述文本特征作为外部引入特征，对所述容器中的物品的申报类别进行甄别。例如，以图像特征作为跨模态解码器的输入信息，以文本特征作为跨模态解码器的注意力机制的外部引入特征，利用跨模态解码器对容器中的物品的申报类别进行甄别。在一个实施例中，特征融合子系统 130 可以用于执行操作 S350。

结论判定子系统 140 可以用于当容器中至少一个物品的申报类别不属于申报信息中的申报类别时，确定申报信息存疑。在一个实施例中，结论判定子系统 140 可以用于执行操作 S360。

图 11 示意性示出了根据本公开实施例的用于检验申报信息真实性的系统中特征提取子系统 120 的框图。

如图 11 所示，根据本公开的一些实施例，特征提取子系统 120 可以包括图像预处理模块 100'、图像特征提取模块 100、以及图像特征映射模块 101。其中，图像特征包括与机检辐射图像中的 N1 个物品的图像信息分别对应的 N1 个第一图像特征向量，其中，N1 为大于或等于 1 的整数。

图像预处理模块 100' 可以用于利用目标检测算法将机检辐射图像中的不同物品划分成独立的图像块，得到 N1 个图像块。在一个实施例中，图像预处理模块 100' 可以执行操作 S331。

图像特征提取模块 100 可以用于提取每个图像块对应的第二图像特征向量。在一个实施例中，图像特征提取模块 100 可以执行操作 S332。

图像特征映射模块 101 可以用于基于每个图像块对应的第二图像特征向量，得到与图像块所表示的物品的图像信息对应的第一图像特征向量。在一个实施例中，图像特征提取模块 101 可以执行操作 S333。

继续参考图 11，根据本公开的另一些实施例，特征提取子系统 120 还可以包括申报信息预处理模块 102' 以及文本特征提取投影模块 102。其中，文本特征可以包括与申报信息中的 N2 种物品分别对应的文本特征向量，其中 N2 为大于或等于 1 的整数。

申报信息预处理模块 102' 用于提取申报信息中每一种物品的名称信息和规格型号信息，并且对于每一种物品，将名称信息处理为第一语句，将规格型号信息处理为第二语句。在一个实施例中，申报信息预处理模块 102' 可以执行操作 S341 和操作 S342。

文本特征提取投影模块 102 用于以对应于同一种物品的第一语句和第二语句作为输入，对该物品所属的申报类别进行分类，并且以文本特征提取投影模块针对每一种物品的类别输出结果作为该物品对应文本特征向量，其中，针对 N2 种物品对应得到 N2 个文本特征向量。在一些实施例中，文本特征提取投影模块 102 可以执行操作 S343 和操作 S344。

图 12 示意性示出了根据本公开另一实施例的用于检验申报信息真实性的系统的整体结构示意图。

如图 12 所示，根据该实施例的用于检验申报信息真实性的系统 1200 可以包括特征提取子系统 120 和跨模态解码器 200。系统 1200 会输出整体结论 3 和分块结论 4。特征提取子系统 120 可以参考图 13 所示，跨模态解码器 200 可以参考图 14 所示。

图 13 示意性示出了根据本公开一实施例的用于检验申报信息真实性的系统 1200 中特征提取子系统 120 的结构示意。

如图 13 所示, 根据该实施例的特征提取子系统 120 可以包括图像特征提取和申报信息提取两个分支。

其中, 图像特征提取分支包括图像特征提取模块 100 和图像特征映射模块 101, 输出图像特征 1。

申报信息提取分支包括文本特征提取投影模块 102, 输出文本特征 2。

图 14 示意性示出了根据本公开一实施例的用于检验申报信息真实性的系统 1200 中跨模态解码器 200 的结构示意。

如图 14 所示, 跨模态解码器 200 由多头自注意力模块 201、特征加和与归一化模块 202、多头互注意力模块 203 以及前馈网络 204 组合而成。输入为由 120 提取得到的图像特征 1 和文本特征 2。其中, 图像特征 1 作为多头自注意力模块 201 的主要输入信息, 文本特征 2 作为增强信息引入到多头互注意力模块 203 中。

结合图 12~图 14, 根据本公开的实施例, 系统 1200 在工作时, 对于机检辐射图像, 利用目标检测法获取图像中不同物品(以图像的边界和纹理对图像中的物品块进行区分)的坐标位置, 并根据坐标位置提取图像块, 作为图像特征提取模块 100 的输入。

其中, 图像块在进入图像特征提取模块 100 之前, 可以进行数据增强处理, 包括旋转, `resize` 到固定大小、去均值、标准化等一系列操作;

当将固定大小的图像块输入到图像特征提取模块 100 中时, 利用交叉熵进行约束。为了能够获取较好的图像特征, 可以采用 `resnet50+SE-block` 组合的形式, 对图像块的特征进行提取, 抽取网络的倒数第二层(即, 全连接层 `fc1` 的输出结果)作为图像块的特征(即, 第二图像特征向量)。

然后将提取到的图像块的特征带入到图像特征映射模块 101 中。图像特征映射模块 101 的结构采用 `transformer` 编码器结构, 可以与图像块对应的位置坐标处理为向量, 然后将图像块的特征与图像块对应的位置坐标(沿着特征维度的方向连接起来)连接起来, 输入到 `transformer` 编码器中, 得到新的具有空间位置关系的图像块特征(即, 第一图像特征向量)。

对于申报单中的物品申报信息, 可以先进行如下步骤的预处理操作: 去除规格型号中的独立数字, 删除部分停用词及符号, 和/或统一文本中的英文字母为小写。

然后，将处理好的物品的名称信息作为句子一，将申报的物品的规格型号作为句子二，输入到文本特征提取投影模块 102(例如，BERT 模型)进行文本特征提取。在文本特征提取投影模块 102 的训练过程中，可以采用多元交叉熵作为损失函数，对训练过程进行约束，将模型中最后一层的池化结果作为该条申报物品的类别输出。其中，当一个申报单中一个商品编号下包括多品种的物品时，需要将申报的物品名称和规格型号逐一对应，作为多种物品进行处理，重复对单个物品的申报信息的处理过程，通过 BERT 模型的前馈过程，得到每种物品的申报信息的文本特征，并将其连接起来。

然后，将特征提取子系统 120 输出的图像特征 1 和文本特征 2 代入到跨模态解码器 200 中，跨模态解码器 200 可以采用的 transformer 解码器模块，将图像特征 1 作为跨模态解码器 200 的 value 输入，将文本特征 2 引入互注意力模块 203 中，通过特征之间的交互操作，完成跨模态的信息融合。每个图像块的图像特征与申报信息中的文本特征交互后，都会得到一个对应的结果，该结果为多类别的概率输出，通过该值判断容器内装载的货物与申报信息中的申报类别是否一致。

根据本公开的实施例的模块、子模块、单元、子单元中的任意多个、或其中任意多个的至少部分功能可以在一个模块中实现。根据本公开实施例的模块、子模块、单元、子单元中的任意一个或多个可以被拆分成多个模块来实现。根据本公开实施例的模块、子模块、单元、子单元中的任意一个或多个可以至少被部分地实现为硬件电路，例如现场可编程门阵列(FPGA)、可编程逻辑阵列(PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(ASIC)，或可以通过对电路进行集成或封装的任何其他的合理方式的硬件或固件来实现，或以软件、硬件以及固件三种实现方式中任意一种或以其中任意几种的适当组合来实现。或者，根据本公开实施例的模块、子模块、单元、子单元中的一个或多个可以至少被部分地实现为计算机程序模块，当该计算机程序模块被运行时，可以执行相应的功能。

例如，信息获取子系统 110、特征提取子系统 120、特征融合子系统 130、结论判定子系统 140、图像预处理模块 100'、图像特征提取模块 100、

图像特征映射模块 101、申报信息预处理模块 102'、文本特征提取投影模块 102 或者跨模态解码器 200 中的任意多个可以合并在一个模块中实现，或者其中的任意一个模块可以被拆分成多个模块。或者，这些模块中的一个或多个模块的至少部分功能可以与其他模块的至少部分功能相结合，并在一个模块中实现。根据本公开的实施例，信息获取子系统 110、特征提取子系统 120、特征融合子系统 130、结论判定子系统 140、图像预处理模块 100'、图像特征提取模块 100、图像特征映射模块 101、申报信息预处理模块 102'、文本特征提取投影模块 102 或者跨模态解码器 200 中的至少一个可以至少被部分地实现为硬件电路，例如现场可编程门阵列(FPGA)、可编程逻辑阵列(PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(ASIC)，或可以通过对电路进行集成或封装的任何其他的合理方式等硬件或固件来实现，或以软件、硬件以及固件三种实现方式中任意一种或以其中任意几种的适当组合来实现。或者，信息获取子系统 110、特征提取子系统 120、特征融合子系统 130、结论判定子系统 140、图像预处理模块 100'、图像特征提取模块 100、图像特征映射模块 101、申报信息预处理模块 102'、文本特征提取投影模块 102 或者跨模态解码器 200 中的至少一个可以至少被部分地实现为计算机程序模块，当该计算机程序模块被运行时，可以执行相应的功能。

图 15 示意性示出了适于实现根据本公开实施例的用于检验申报信息真实性的方法的电子设备的框图。图 15 示出的电子设备 1500 仅仅是一个示例，不应对本公开实施例的功能和使用范围带来任何限制。

如图 15 所示，根据本公开实施例的电子设备 1500 包括处理器 1501，其可以根据存储在只读存储器(ROM)1502 中的程序或者从存储部分 1508 加载到随机访问存储器(RAM) 1503 中的程序而执行各种适当的动作和处理。处理器 1501 例如可以包括通用微处理器(例如 CPU)、指令集处理器和/或相关芯片组和/或专用微处理器(例如，专用集成电路(ASIC))，等等。处理器 1501 还可以包括用于缓存用途的板载存储器。处理器 1501 可以包括用于执行根据本公开实施例的方法流程的不同动作的单一处理单元或者是多个处理单元。

在 RAM 1503 中, 存储有电子设备 1500 操作所需的各种程序和数据。处理器 1501、ROM 1502 以及 RAM 1503 通过总线 1504 彼此相连。处理器 1501 通过执行 ROM 1502 和/或 RAM 1503 中的程序来执行根据本公开实施例的方法流程的各种操作。需要注意, 所述程序也可以存储在除 ROM 1502 和 RAM 1503 以外的一个或多个存储器中。处理器 1501 也可以通过执行存储在所述一个或多个存储器中的程序来执行根据本公开实施例的方法流程的各种操作。

根据本公开的实施例, 电子设备 1500 还可以包括输入/输出 (I/O) 接口 1505, 输入/输出 (I/O) 接口 1505 也连接至总线 1504。电子设备 1500 还可以包括连接至 I/O 接口 1505 的以下部件中的一项或多项: 包括键盘、鼠标等的输入部分 1506; 包括诸如阴极射线管 (CRT)、液晶显示器 (LCD) 等以及扬声器等的输出部分 1507; 包括硬盘等的存储部分 1508; 以及包括诸如 LAN 卡、调制解调器等网络接口卡的通信部分 1509。通信部分 1509 经由诸如因特网的网络执行通信处理。驱动器 1510 也根据需要连接至 I/O 接口 1505。可拆卸介质 1511, 诸如磁盘、光盘、磁光盘、半导体存储器等等, 根据需要安装在驱动器 1510 上, 以便于从其上读出的计算机程序根据需要被安装入存储部分 1508。

根据本公开的实施例, 根据本公开实施例的方法流程可以被实现为计算机软件程序。例如, 本公开的实施例包括一种计算机程序产品, 其包括承载在计算机可读存储介质上的计算机程序, 该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中, 该计算机程序可以通过通信部分 1509 从网络上被下载和安装, 和/或从可拆卸介质 1511 被安装。在该计算机程序被处理器 1501 执行时, 执行本公开实施例的系统中限定的上述功能。根据本公开的实施例, 上文描述的系统、设备、装置、模块、单元等可以通过计算机程序模块来实现。

本公开还提供了一种计算机可读存储介质, 该计算机可读存储介质可以是上述实施例中描述的设备/装置/系统中所包含的; 也可以是单独存在, 而未装配入该设备/装置/系统中。上述计算机可读存储介质承载有一个或者多个程序, 当上述一个或者多个程序被执行时, 实现根据本公开实施例的方法。

根据本公开的实施例，计算机可读存储介质可以是非易失性的计算机可读存储介质，例如可以包括但不限于：便携式计算机磁盘、硬盘、随机访问存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、便携式紧凑磁盘只读存储器（CD-ROM）、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。例如，根据本公开的实施例，计算机可读存储介质可以包括上文描述的 ROM 1502 和/或 RAM 1503 和/或 ROM 1502 和 RAM 1503 以外的一个或多个存储器。

本公开的实施例还包括一种计算机程序产品，其包括计算机程序，该计算机程序包含用于执行本公开实施例所提供的方法的程序代码，当计算机程序产品在电子设备上运行时，该程序代码用于使电子设备实现本公开实施例所提供的图像识别方法。

在该计算机程序被处理器 1501 执行时，执行本公开实施例的系统/装置中限定的上述功能。根据本公开的实施例，上文描述的系统、装置、模块、单元等可以通过计算机程序模块来实现。

在一种实施例中，该计算机程序可以依托于光存储器件、磁存储器件等有形存储介质。在另一种实施例中，该计算机程序也可以在网络介质上以信号的形式进行传输、分发，并通过通信部分 1509 被下载和安装，和/或从可拆卸介质 1511 被安装。该计算机程序包含的程序代码可以用任何适当的网络介质传输，包括但不限于：无线、有线等等，或者上述的任意合适的组合。

根据本公开的实施例，可以以一种或多种程序设计语言的任意组合来编写用于执行本公开实施例提供的计算机程序的程序代码，具体地，可以利用高级过程和/或面向对象的编程语言、和/或汇编/机器语言来实施这些计算程序。程序设计语言包括但不限于诸如 Java，C++，python，“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、部分在远程计算设备上执行、或者完全在远程计算设备或服务器上执行。在涉及远程计算设备的情形中，远程计算设备可以通过任意种类的网络，包括局域网（LAN）或广域网（WAN），连接

到用户计算设备，或者，可以连接到外部计算设备（例如利用因特网服务提供商来通过因特网连接）。

附图中的流程图和框图，图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图或流程图中的每个方框、以及框图或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

以上对本公开的实施例进行了描述。但是，这些实施例仅仅是为了说明的目的，而并非为了限制本公开的范围。尽管在以上分别描述了各实施例，但是这并不意味着各个实施例中的措施不能有利地结合使用。本公开的范围由所附权利要求及其等同物限定。不脱离本公开的范围，本领域技术人员可以做出多种替代和修改，这些替代和修改都应落在本公开的范围之内。

权 利 要 求

1. 一种用于检验申报信息真实性的方法，包括：

获取对装载了物品的容器进行扫描而获得的机检辐射图像；

获取对所述容器内的物品进行申报的申报信息；

对所述机检辐射图像中的物品的图像信息进行识别，得到与所述机检辐射图像对应的图像特征；

对所述申报信息中的物品的文本信息进行识别，得到与所述申报信息对应的文本特征；其中，所述文本特征用于表征所述申报信息中的物品所属的申报类别；

以所述图像特征作为输入信息，以所述文本特征作为外部引入特征，对所述容器中的物品的申报类别进行甄别；以及

当所述容器中至少一个物品的申报类别不属于所述申报信息中的申报类别时，确定所述申报信息存疑。

2. 根据权利要求 1 所述的方法，其中，所述图像特征包括与所述机检辐射图像中的 N1 个物品的图像信息分别对应的 N1 个第一图像特征向量，其中，N1 为大于或等于 1 的整数。

3. 根据权利要求 2 所述的方法，其中，所述对所述机检辐射图像中的物品的图像信息进行识别，得到与所述机检辐射图像对应的图像特征包括：

利用目标检测算法将所述机检辐射图像中的不同物品划分成独立的图像块，得到 N1 个图像块；

提取每个所述图像块对应的第二图像特征向量；以及

基于每个所述图像块对应的所述第二图像特征向量，得到与所述图像块所表示的物品的图像信息对应的所述第一图像特征向量。

4. 根据权利要求 3 所述的方法，其中，所述提取每个所述图像块对应的第二图像特征向量包括：

利用图像特征提取模块对每个所述图像块进行图像识别，得到每个所述图像块对应的所述第二图像特征向量；

其中，所述图像特征提取模块包括卷积神经网络。

5. 根据权利要求 4 所述的方法，其中，所述图像特征提取模块包括以 resnet 作为基础网络，在 resnet 池化层后添加 SE-block 的网络结构。

6. 根据权利要求 3 所述的方法，其中，所述基于每个所述图像块对应的所述第二图像特征向量，得到与所述图像块所表示的物品的图像信息对应的所述第一图像特征向量包括：

获取每个所述图像块在所述机检辐射图像中的位置信息；

基于同一个所述图像块对应的所述第二图像特征向量和所述位置信息，得到所述第一图像特征向量。

7. 根据权利要求 6 所述的方法，其中，所述基于同一个所述图像块对应的所述第二图像特征向量和所述位置信息，得到所述第一图像特征向量包括：

利用所述图像块的所述位置信息处理所述第二图像特征向量；

将处理后的所述第二图像特征向量输入编码器；以及

获得所述编码器的输出，以得到与所述图像块对应的所述第一图像特征向量。

8. 根据权利要求 7 所述的方法，其中，所述以所述图像特征作为输入信息，以所述文本特征作为外部引入特征，对所述容器中的物品的申报类别进行甄别，包括：

以所述图像特征作为跨模态解码器的输入信息，以所述文本特征作为所述跨模态解码器的注意力机制的外部引入特征，利用所述跨模态解码器对所述容器中的物品的申报类别进行甄别。

9. 根据权利要求 8 所述的方法，其中，所述编码器与所述跨模态解码器联合训练。

10. 根据权利要求 9 所述的方法，其中，所述编码器采用 transformer 编码器模型。

11. 根据权利要求 10 所述的方法，其中，所述跨模态解码器采用 transformer 解码器模型。

12. 根据权利要求 1~11 任意一项所述的方法, 其中, 所述文本特征包括与所述申报信息中的 $N2$ 种物品分别对应的文本特征向量, 其中 $N2$ 为大于或等于 1 的整数。

13. 根据权利要求 12 所述的方法, 其中, 所述对所述申报信息中的物品的文本信息进行识别, 得到与所述申报信息对应的文本特征包括:

提取所述申报信息中每一种物品的名称信息和规格型号信息;

对于每一种物品, 将所述名称信息处理为第一语句, 将所述规格型号信息处理为第二语句;

将对应于同一种物品的所述第一语句和所述第二语句作为文本特征提取投影模块的输入, 利用所述文本特征提取投影模块对该物品所属的申报类别进行分类; 以及

以所述文本特征提取投影模块针对每一种物品的类别输出结果作为该物品对应所述文本特征向量; 其中, 针对 $N2$ 种物品对应得到 $N2$ 个所述文本特征向量。

14. 根据权利要求 13 所述的方法, 其中, 所述文本特征提取投影模块采用 BERT 模型。

15. 一种用于检验申报信息真实性的系统, 包括:

信息获取子系统, 用于获取对装载了物品的容器进行扫描而获得的机检辐射图像, 以及获取对所述容器内的物品进行申报的申报信息;

特征提取子系统, 用于对所述机检辐射图像中的物品的图像信息进行识别, 得到与所述机检辐射图像对应的图像特征; 以及对所述申报信息中的物品的文本信息进行识别, 得到与所述申报信息对应的文本特征; 其中, 所述文本特征用于表征所述申报信息中的物品所属的申报类别;

特征融合子系统, 用于以所述图像特征作为输入信息, 以所述文本特征作为外部引入特征, 对所述容器中的物品的申报类别进行甄别; 以及

结论判定子系统, 用于当所述容器中至少一个物品的申报类别不属于所述申报信息中的申报类别时, 确定所述申报信息存疑。

16. 一种电子设备，包括：

一个或多个存储器，存储有可执行指令；以及

一个或多个处理器，执行所述可执行指令，以实现根据权利要求 1~14 中任一项所述的方法。

17. 一种计算机可读存储介质，其上存储有可执行指令，该指令被处理器执行时使处理器执行根据权利要求 1~14 中任一项所述的方法。