

图像检测方法、装置、电子设备和存储介质

相关申请的交叉引用

本申请要求申请号为 202311108453.1，题为“图像检测方法、装置、电子设备和存储介质”、申请日为 2023 年 8 月 30 日的中国发明专利申请的优先权，通过引用的方式将该申请整本并入本文。

技术领域

本公开涉及计算机技术领域，具体涉及一种图像检测方法、装置、电子设备和存储介质。

背景技术

当前人工智能生成图像中的虚拟人像瑕疵是一个普遍存在的问题，例如手指数量异常、肢体比例异常、肢体关节连接不自然等，这严重影响了人工智能生成图像的质量。对此，目前尚且缺乏一种有效的判别方法，以对人工智能生成图像中虚拟人像是否存在瑕疵进行判别。

发明内容

提供该发明内容部分以便以简要的形式介绍构思，这些构思将在后面的具体实施方式部分被详细描述。该发明内容部分并不旨在标识要求保护的技术方案的关键特征或必要特征，也不旨在用于限制所要求的保护的技术方案的范围。

第一方面，根据本公开的一个或多个实施例，提供了一种图像检测方法，包括：

确定待测图像中是否存在目标对象；

响应于所述待测图像中存在所述目标对象，提取所述待测图像中所述目标对象的结构关键点信息；

基于所述结构关键点信息和所述待测图像，确定所述待测图像中所述目标对象的分类。

第二方面，根据本公开的一个或多个实施例，提供了一种图像检测装置，包括：

确定单元，用于确定待测图像中是否存在目标对象；

提取单元，用于响应于所述待测图像中存在所述目标对象，提取所述待测图像中所述目标对象的结构关键点信息；

分类单元，用于基于所述结构关键点信息和所述待测图像，确定所述待测图像中所述目标对象的分类。

第三方面，根据本公开的一个或多个实施例，提供了一种电子设备，包括：至少一个存储器和至少一个处理器；其中，所述存储器用于存储程序代码，所述处理器用于调用所述存储器所存储的程序代码以使所述电子设备执行根据本公开的一个或多个实施例提供的方法。

第四方面，根据本公开的一个或多个实施例，提供了一种非暂态计算机存储介质，所述非暂态计算机存储介质存储有程序代码，所述程序代码被计算机设备执行时，使得所述计算机设备执行根据本公开的一个或多个实施例提供的方法。

附图说明

结合附图并参考以下具体实施方式，本公开各实施例的上述和其他特征、优点及方面将变得更加明显。贯穿附图中，相同或相似的附图标记表示相同或相似的元素。应当理解附图是示意性的，原件和元素不一定按照比例绘制。

图1为本公开一实施例提供的图像检测方法的流程图；

图2为本公开另一实施例提供的图像检测方法的流程图；

图3为本公开又一实施例提供的图像检测方法的流程图；

图4为本公开一实施例提供的图像检测装置的结构示意图；

图5为根据本公开一实施例提供的电子设备的结构示意图。

具体实施方式

下面将参照附图更详细地描述本公开的实施例。虽然附图中显示了本公开的某些实施例，然而应当理解的是，本公开可以通过各种形式来实现，而且不应该被解释为限于这里阐述的实施例，相反提供这些实施例是为了更加透彻和完整地理解本公开。应当理解的是，本公开的附图及实施例仅用于示例性作用，并非用于限制本公开的保护范围。

应当理解，本公开的实施方式中记载的步骤可以按照不同的顺序执行，和/或并行执行。此外，实施方式可以包括附加的步骤和/或省略执行示出的步骤。本公开的范围在此方面不受限制。

本文使用的术语“包括”及其变形是开放性包括，即“包括但不限于”。术语“基于”是“至少部分地基于”。术语“一个实施例”表示“至少一个实施例”；术语“另一实施例”表示“至少一个另外的实施例”；术语“一些实施例”表示“至少一些实施

例”。术语“响应于”以及有关的术语是指一个信号或事件被另一个信号或事件影响到某个程度，但不一定是完全地或直接地受到影响。如果事件 x “响应于”事件 y 而发生，则 x 可以直接或间接地响应于 y。例如，y 的出现最终可能导致 x 的出现，但可能存在其它中间事件和/或条件。在其它情形中，y 可能不一定导致 x 的出现，并且即使 y 尚未发生，x 也可能发生。此外，术语“响应于”还可以意味着“至少部分地响应于”。

术语“确定”广泛涵盖各种各样的动作，可包括获取、演算、计算、处理、推导、调研、查找（例如，在表、数据库或其他数据结构中查找）、探明、和类似动作，还可包括接收（例如，接收信息）、访问（例如，访问存储器中的数据）和类似动作，以及解析、选择、选取、建立和类似动作等等。其他术语的相关定义将在下文描述中给出。其他术语的相关定义将在下文描述中给出。

可以理解的是，本技术方案所涉及的数据（包括但不限于数据本身、数据的获得或使用）应当遵循相关法律法规的规定。

可以理解的是，在使用本公开各实施例公开的技术方案之前，均应当根据相关法律法规通过适当的方式对本公开所涉及个人信息类型、使用范围、使用场景等告知用户并获得用户的授权。例如，在响应于接收到用户的主动请求时，向用户发送提示信息，以明确地提示用户，其请求执行的操作将需要获得和使用到用户的个人信息，从而使得用户可以根据提示信息来自主地选择是否向执行本公开技术方案的操作的电子设备、应用程序、服务器或存储介质等软件或硬件提供个人信息。

作为一种可选的但非限制性的实现方式，响应于接收到用户的主动请求，向用户发送提示信息的方式，例如可以是弹窗的方式，弹窗中可以以文字的方式呈现提示信息。此外，弹窗中还可以承载供用户选择“同意”或“不同意”向电子设备提供个人信息的选择控件。

可以理解的是，对于根据本公开各实施例提供的方法所生成的图像，应当遵循相关法律法规的规定对其进行处理。例如，根据规定采取技术措施添加不影响用户使用的标识，或者，根据规定在合理位置、区域进行显著标识，向公众提示深度合成情况。

可以理解的是，上述通知和获得用户授权过程、以及对图像的处理，仅是示意性的，不对本公开的实现方式构成限定，其他满足相关法律法规的方式也可应用于本公开的实现方式中。

需要注意，本公开中提及的“第一”、“第二”等概念仅用于对不同的装置、模块或单元进行区分，并非用于限定这些装置、模块或单元所执行的功能的顺序或者相互依存关系。

需要注意，本公开中提及的“一个”、“多个”的修饰是示意性而非限制性的，本领域技术人员应当理解，除非在上下文另有明确指出，否则应该理解为“一个或多个”。

为了本公开的目的，短语“A和/或B”意为(A)、(B)或(A和B)。

本公开实施方式中的多个装置之间所交互的消息或者信息的名称仅用于说明性的目的，而并不是用于对这些消息或信息的范围进行限制。

针对人工智能生成内容中虚拟人像的肢体瑕疵判别，当前缺乏有针对性的方案，相关技术普遍将其视为一个分类问题，即，通过样本图像及对应的人工标注的、指示肢体有无瑕疵的标签训练分类模型，并利用该分类模型对图像进行判别。但由于判别肢体瑕疵本身较为困难，肢体在不同视角下、不同人体动作下对应的正常状态各不相同，这就要求分类模型能够深入理解肢体的状态。而相关技术提供的仅基于样本和人工标注的标签来训练模型的方案，易导致在训练过程中，分类模型容易过拟合到图像中其他易拟合的特征上去（如画面整体复杂程度），使得训练得到的分类模型并未真正理解肢体瑕疵。此外，由于人工标记样本数据时难免会存在误标注的情况，这种数据噪声的存在进一步加剧了模型在训练过程中拟合到其他特征的概率。

针对上述问题，本公开提供了一种图像检测方法，通过基于待测图像以及从待测图像中提取的目标对象的结构关键点信息对目标对象进行分类，从而能够使分类模型在训练和判别过程中更加关注目标对象，能够避免模型过拟合到其他特征上，进而增强模型对数据标注噪声的鲁棒性，提升对目标对象分类的准确度。

参考图1，图1示出了本公开一实施例提供的图像检测方法100的流程图，方法100包括步骤S110-步骤S130。

步骤S110：确定待测图像中是否存在目标对象；

步骤S120：响应于所述待测图像中存在所述目标对象，提取所述待测图像中所述目标对象的结构关键点信息；

步骤S130：基于所述结构关键点信息和所述待测图像，确定所述待测图像中所述目标对象的分类。

目标对象的结构关键点用于决定目标对象的结构形态。在一个具体实施方式中，目标对象可以包括虚拟人像，虚拟人像例如可以包括人工智能生成内容中包含的人像；相应地，目标对象的结构关键点信息可以包括虚拟人体关键点信息，例如，可以包括鼻、眼、耳、肩、手肘、手腕、臀、膝、脚踝中的部分或全部关键点的信息。

在一些实施例中，步骤S130可以用于确定所述待测图像中是否存在虚拟人像瑕疵。下面以判断虚拟人像是否存在瑕疵的分类场景进行示例性说明。

参考图 2，其示出了一种判断图像中虚拟人像是否存在瑕疵的方法 200，包括步骤 S201-步骤 S204。

在步骤 201 中，获取待测图像。示例性地，待测图像可以包括人工智能生成内容或其他虚拟图像。

在步骤 202 中，对待测图像进行人体检测，以确定是否存在虚拟人像。在一些实施例中，可以通过目标检测算法确定待测图像中是否存在虚拟人像。目标检测的任务是找出图像中所有感兴趣的目标（物体），确定他们的类别和位置。目标检测算法可以采用诸如基于区域的卷积神经网络的目标检测算法、单步多框目标检测算法、基于视网膜网络的目标检测算法等，但本公开不限于此。

若确定不存在虚拟人像，则可以直接确定待测图像不存在虚拟人像瑕疵，若确定存在虚拟人像，则执行步骤 203。

在步骤 203 中，提取待测图像的虚拟人体关键点信息。在一些实施例中，可以通过人体关键点估计算法提取待测图像的虚拟人体关键点信息，人体关键点估计算法可以采用诸如基于高分辨网络、卷积姿态机、堆叠沙漏网络、级联金字塔网络的人体姿态估计方法，但本公开不限于此。在一个具体实施方式中，可以提取待测图像中虚拟人像的双手、双肘、双脚、双腿所对应的 8 个坐标，作为虚拟人体关键点信息。

在步骤 204 中，可以将虚拟人体关键点信息和待测图像一起作为训练好的分类模型的输入，以确定待测图像中的虚拟人像是否存在瑕疵。在一个具体实施方式中，分类模型被训练为能够基于输入的图像和虚拟人体关键点信息确定图像中虚拟人像存在瑕疵的置信度。示例性，可以将样本图像以及基于样本图像提取的虚拟人体关键点信息作为模型的输入，以样本图像对应的人工标注的标签（例如关于样本图像是否存在虚拟人像瑕疵）作为期望输出，训练分类模型。

人工标注的标签可以关于样本图像是否存在虚拟人像瑕疵，还可以根据其他应用需要指代其他分类类型，本公开在此不做限制。在一个具体实施中，正标签可以为 1，负标签可以为 0。

需要说明的是，根据本公开提供的图像检测方法，除可以用于判断外虚拟人像是否存在瑕疵的分类场景外，还能用于判别动物、机器人、虚拟形象等具有特定结构的目标对象的分类，本公开在此不做限制。

根据本公开的一个或多个实施例，通过基于待测图像以及从待测图像中提取的目标对象的结构关键点信息对目标对象进行分类，从而能够使分类模型在训练和判别过程中更加关注目标对象，能够避免模型过拟合到其他特征上，进而增强模型对数据标注噪声

的鲁棒性，提升对目标对象分类的准确度。

在一些实施例中，步骤 S130 包括：

步骤 A1：基于所述待测图像生成第一目标特征，所述第一目标特征包括与所述目标对象关联的特征；

步骤 A2：对所述结构关键点信息进行特征转换得到第二目标特征；

步骤 A3：基于所述第一目标特征和所述第二目标特征生成第三目标特征；

步骤 A4：基于所述第三目标特征确定所述待测图像中所述目标对象的分类。

在一些实施例中，所述第二目标特征对应的通道数与所述第一目标特征对应的通道数相同，通过使第二目标特征对应的通道数与所述第一目标特征对应的通道数相同，可以便于后续连接第一目标特征和第二目标特征以得到第三目标特征。

在一些实施例中，在所述基于所述第三目标特征确定所述待测图像中所述目标对象的分类之前，对所述第三目标特征进行语义单元融合处理，以使所述第三目标特征中的语义单元能够融合所述第三目标特征中其他语义单元的信息，即，使语义单元之间能够互相融合。其中，语义单元是具有语义信息的单元，例如令牌（token）。进一步地，后续可以基于语义单元融合处理后的第三目标特征，确定所述待测图像中所述目标对象的分类。

在本实施例中，通过对所述第三目标特征进行语义单元融合处理，可以进一步在语义单元这一层级上，将来自待测图像的特征和来自结构关键点信息的特征进行一定程度的融合，从而能够进一步提升分类模型对图像中目标对象部分的区域的关注。

在一些具体实施方式中，可以利用转换网络中的自注意力机制基于各语义单元之间的相关性确定融合权重以实现语义单元融合，或在卷积神经网络中通过扩大卷积核的尺寸实现语义单元融合，或在多层感知机中通过令牌之间的全连接实现语义单元融合，但本公开不限于此。

需要说明的是，本实施例中的语义单元融合处理，可以是全局范围的语义单元融合，还可以针对部分语义单元的融合，本公开在此不做限制。

在一些实施例中，步骤 A4 包括：

步骤 a41：将所述第三目标特征输入分类网络，得到目标对象存在瑕疵置信度和/或目标对象不存在瑕疵置信度；

步骤 a42：基于所述目标对象存在瑕疵置信度和/或所述目标对象不存在瑕疵置信度，得到目标对象的类别。

在一些实施例中，将第三目标特征输入分类网络之前，还可以对第三目标特征进行

其他处理，例如语义单元融合处理、预设维度求均值处理等，本公开在此不做限制。

参见图 3，下面以判断虚拟人像是否存在瑕疵的分类场景进行示例性说明。

第一步，基于待测图像的虚拟人体关键点信息可以获得一个归一化张量 T_i ，其可以表示为 (B, N, C) ，其中， B 代表训练样本的批大小，即单次传递给程序用以训练的数据（样本）个数； N 表示语义单元的数量； C 表示通道数。示例性地，假设对于每个待测图像获取其中 8 个人体关键点在 x 和 y 这两个方向上的坐标 (x, y) ，则 N 可以为 8， C 可以为 2。

第二步，将张量 T_i 输入特征提取网络进行特征转换得到张量 T_f ，其可以表示为 (B, N_1, C_1) 。在一个具体实施方式中，特征提取网络可以为转换网络，其是一种基于自注意力机制的神经网络模型，用于处理序列数据。

第三步，将待测图像（例如 RGB 图像）输入骨干网络提取特征，得到张量 T_m ，其可以表示为 (B, N_2, C_1) 。训练良好的骨干网络被期望用于从待测图像中提取与分类任务关联的特征，在关于虚拟人像的分类场景下，用于提取关于虚拟人像的结构、数量的特征。当然，在实际应用中，骨干网络提取的特征还可能包括不被期望提取的特征，例如图像中与虚拟人像无关的建筑物或其他物体的特征。在一个具体实施方式中，骨干网络可以采用视觉转换网络或者其他图像特征提取网络，但本公开不限于此。

第四步，将张量 T_f 和张量 T_m 在语义单元的维度上进行连接，得到张量 T_c ，其可以表示为 (B, N_1+N_2, C_1) 。示例性地，连接操作可以通过连接函数（Concat）实现，其可以在指定维度上对特征进行连接。

第五步，其张量 T_c 输入特征融合网络，得到张量 T_o ，其可以表示为 (B, N_3, C_2) 。特征融合网络可以用于对输入的特征进行语义单元融合处理，以使特征中的各语义单元能够基于其与其他语义单元之间的相关性而进行语义单元融合，从而提升信息交互程度。示例性地，特征融合网络可以为转换网络。

第六步，在语义单元的维度上（即 (B, N, C) 中的“ N ”）对张量 T_o 进行求均值操作，并输入分类头网络，得到最终关于虚拟人像是否存在瑕疵的分类，分类结果可以表示为“虚拟人像存在瑕疵”的置信度和/或“虚拟人像不存在瑕疵”的置信度。示例性地，分类头网络可以为一层全连接层。在本步骤中，通过在语义单元的维度上对张量 T_o 进行求均值操作，能够易于采用分类头网络对其进行处理。

在本实施例中，分类网络包括特征提取网络、骨干网络、特征融合网络和分类头网络。示例性地，分类网络训练时，可以将样本图像及其对应的虚拟人体关键点信息分别输入骨干网络和特征提取网络，并依次经过特征融合网络和分类头网络得到预测结果，并基于预测结果与样本图像对应的标签的损失（例如交叉熵损失）调整特征提取网络、

骨干网络、特征融合网络和分类头网络的参数，以使损失函数收敛。在一个具体实施方式中，可以通过预训练参数对骨干网络进行初始化，以加速收敛。

相应地，参考图 4，根据本公开一实施例提供了一种图像检测装置 400，包括：

确定单元 401，用于确定待测图像中是否存在目标对象；

提取单元 402，用于响应于所述待测图像中存在所述目标对象，提取所述待测图像中所述目标对象的结构关键点信息；

分类单元 403，用于基于所述结构关键点信息和所述待测图像，确定所述待测图像中所述目标对象的分类。

在一些实施例中，所述目标对象包括虚拟人像，所述目标对象的结构关键点信息包括虚拟人体关键点信息。

在一些实施例中，分类单元包括：

特征提取子单元，用于基于所述待测图像生成第一目标特征，所述第一目标特征包括与所述目标对象关联的特征；

特征转换子单元，用于对所述结构关键点信息进行特征转换得到第二目标特征；

特征生成子单元，用于基于所述第一目标特征和所述第二目标特征生成第三目标特征；以及

分类子单元，用于基于所述第三目标特征确定所述待测图像中所述目标对象的分类。

在一些实施例中，所述分类子单元用于将所述第三目标特征输入分类网络，得到目标对象存在瑕疵置信度和/或目标对象不存在瑕疵置信度，基于所述目标对象存在瑕疵置信度和/或所述目标对象不存在瑕疵置信度，得到目标对象的类别。

在一些实施例中，图像检测装置还包括：

特征融合子单元，用于在所述基于所述第三目标特征确定所述待测图像中所述目标对象的分类之前，对所述第三目标特征进行语义单元融合处理，以使所述第三目标特征中的语义单元能够融合所述第三目标特征中其他语义单元的信息；

所述分类子单元用于基于语义单元融合处理后的第三目标特征，确定所述待测图像中所述目标对象的分类。

在一些实施例中，所述确定单元还用于响应于所述待测图像中不存在所述目标对象，确定所述待测图像不存在目标对象瑕疵。

对于装置的实施例而言，由于其基本对应于方法实施例，所以相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的，其中作为分离模块说明的模块可以是或者也可以不是分开的。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。本领域普通技术人员在不付出创造性劳动的情况下，即可以理解并实施。

相应地，根据本公开的一个或多个实施例，提供了一种电子设备，包括：

至少一个存储器和至少一个处理器；

其中，存储器用于存储程序代码，处理器用于调用存储器所存储的程序代码以使所述电子设备执行根据本公开一个或多个实施例提供的图像检测方法。

相应地，根据本公开的一个或多个实施例，提供了一种非暂态计算机存储介质，非暂态计算机存储介质存储有程序代码，程序代码可被计算机设备执行来使得所述计算机设备执行根据本公开一个或多个实施例提供的图像检测方法。

下面参考图 5，其示出了适于用来实现本公开实施例的电子设备（例如终端设备或服务器）800 的结构示意图。本公开实施例中的终端设备可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、PDA（个人数字助理）、PAD（平板电脑）、PMP（便携式多媒体播放器）、车载终端（例如车载导航终端）等等的移动终端以及诸如数字 TV、台式计算机等等的固定终端。图 5 示出的电子设备仅仅是一个示例，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 5 所示，电子设备 800 可以包括处理装置（例如中央处理器、图形处理器等）801，其可以根据存储在只读存储器（ROM）802 中的程序或者从存储装置 808 加载到随机访问存储器（RAM）803 中的程序而执行各种适当的动作和处理。在 RAM803 中，还存储有电子设备 800 操作所需的各种程序和数据。处理装置 801、ROM 802 以及 RAM 803 通过总线 804 彼此相连。输入/输出（I/O）接口 805 也连接至总线 804。

通常，以下装置可以连接至 I/O 接口 805：包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置 806；包括例如液晶显示器（LCD）、扬声器、振动器等等的输出装置 807；包括例如磁带、硬盘等的存储装置 808；以及通信装置 809。通信装置 809 可以允许电子设备 800 与其他设备进行无线或有线通信以交换数据。虽然图 5 示出了具有各种装置的电子设备 800，但是应理解的是，并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。

特别地，根据本公开的实施例，上文参考流程图描述的过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装置 809 从网络上被下载和安装，

或者从存储装置 808 被安装，或者从 ROM 802 被安装。在该计算机程序被处理装置 801 执行时，执行本公开实施例的方法中限定的上述功能。

需要说明的是，本公开上述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、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）、光纤、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。而在本公开中，计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：电线、光缆、RF（射频）等等，或者上述的任意合适的组合。

在一些实施方式中，客户端、服务器可以利用诸如 HTTP（HyperText Transfer Protocol，超文本传输协议）之类的任何当前已知或未来研发的网络协议进行通信，并且可以与任意形式或介质的数字数据通信（例如，通信网络）互连。通信网络的示例包括局域网（“LAN”），广域网（“WAN”），网际网（例如，互联网）以及端对端网络（例如，ad hoc 端对端网络），以及任何当前已知或未来研发的网络。

上述计算机可读介质可以是上述电子设备中所包含的；也可以是单独存在，而未装配入该电子设备中。

上述计算机可读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该电子设备执行时，使得该电子设备执行上述的本公开的方法。

可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的计算机程序代码，上述程序设计语言包括面向对象的程序设计语言—诸如 Java、Smalltalk、C++，还包括常规的过程式程序设计语言—诸如“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机上执行、部分地在用户计算机上执行、作为一个独

立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)—连接到用户计算机，或者，可以连接到外部计算机（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，该模块、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，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。

描述于本公开实施例中所涉及到的单元可以通过软件的方式实现，也可以通过硬件的方式来实现。其中，单元的名称在某种情况下并不构成对该单元本身的限定。

本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例如，非限制性地，可以使用的示范类型的硬件逻辑部件包括：现场可编程门阵列（FPGA）、专用集成电路（ASIC）、专用标准产品（ASSP）、片上系统（SOC）、复杂可编程逻辑设备（CPLD）等等。

在本公开的上下文中，机器可读介质可以是有形的介质，其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备，或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦除可编程只读存储器（EPROM 或快闪存储器）、光纤、便捷式紧凑盘只读存储器（CD-ROM）、光学储存设备、磁储存设备、或上述内容的任何合适组合。

根据本公开的一个或多个实施例，提供了一种图像检测方法，包括：确定待测图像中是否存在目标对象；响应于所述待测图像中存在所述目标对象，提取所述待测图像中所述目标对象的结构关键点信息；基于所述结构关键点信息和所述待测图像，确定所述待测图像中所述目标对象的分类。

根据本公开的一个或多个实施例，所述目标对象包括虚拟人像，所述目标对象的结构关键点信息包括虚拟人体关键点信息。

根据本公开的一个或多个实施例，基于所述结构关键点信息和所述待测图像，确定所述待测图像中所述目标对象的分类，包括：基于所述待测图像生成第一目标特征，所述第一目标特征包括与所述目标对象关联的特征；对所述结构关键点信息进行特征转换得到第二目标特征；基于所述第一目标特征和所述第二目标特征生成第三目标特征；基于所述第三目标特征确定所述待测图像中所述目标对象的分类。

根据本公开的一个或多个实施例，所述基于所述第三目标特征确定所述待测图像中所述目标对象的分类，包括：将所述第三目标特征输入分类网络，得到目标对象存在瑕疵置信度和/或目标对象不存在瑕疵置信度；基于所述目标对象存在瑕疵置信度和/或所述目标对象不存在瑕疵置信度，得到所述目标对象的类别。

根据本公开的一个或多个实施例，所述方法还包括：在所述基于所述第三目标特征确定所述待测图像中所述目标对象的分类之前，对所述第三目标特征进行语义单元融合处理，以使所述第三目标特征中的语义单元能够融合所述第三目标特征中其他语义单元的信息；所述基于所述第三目标特征确定所述待测图像中所述目标对象的分类，包括：基于语义单元融合处理后的第三目标特征，确定所述待测图像中所述目标对象的分类。

根据本公开的一个或多个实施例提供的方法，还包括：响应于所述待测图像中不存在所述目标对象，确定所述待测图像不存在目标对象瑕疵。

根据本公开的一个或多个实施例，提供了一种图像检测装置，包括：确定单元，用于确定待测图像中是否存在目标对象；提取单元，用于响应于所述待测图像中存在所述目标对象，提取所述待测图像中所述目标对象的结构关键点信息；分类单元，用于基于所述结构关键点信息和所述待测图像，确定所述待测图像中所述目标对象的分类。

根据本公开的一个或多个实施例，提供了一种电子设备，包括：至少一个存储器和至少一个处理器；其中，所述存储器用于存储程序代码，所述处理器用于调用所述存储器所存储的程序代码以使所述电子设备执行根据本公开的一个或多个实施例提供的图像检测方法。

根据本公开的一个或多个实施例，提供了一种非暂态计算机存储介质，所述非暂态计算机存储介质存储有程序代码，所述程序代码被计算机设备执行时，使得所述计算机设备执行根据本公开的一个或多个实施例提供的图像检测方法。

以上描述仅为本公开的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解，本公开中所涉及的公开范围，并不限于上述技术特征的特定组合而成

的技术方案，同时也应涵盖在不脱离上述公开构思的情况下，由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本公开中公开的（但不限于）具有类似功能的技术特征进行互相替换而形成的技术方案。

此外，虽然采用特定次序描绘了各操作，但是这不应当理解为要求这些操作以所示出的特定次序或以顺序次序执行来执行。在一定环境下，多任务和并行处理可能是有利的。同样地，虽然在上面论述中包含了若干具体实现细节，但是这些不应当被解释为对本公开的范围的限制。在单独的实施例的上下文中描述的某些特征还可以组合地实现在单个实施例中。相反地，在单个实施例的上下文中描述的各种特征也可以单独地或以任何合适的子组合的方式实现在多个实施例中。

尽管已经采用特定于结构特征和/或方法逻辑动作的语言描述了本主题，但是应当理解所附权利要求书中所限定的主题未必局限于上面描述的特定特征或动作。相反，上面所描述的特定特征和动作仅仅是实现权利要求书的示例形式。

权利要求书

1. 一种图像检测方法，包括：

确定待测图像中是否存在目标对象；

响应于所述待测图像中存在所述目标对象，提取所述待测图像中所述目标对象的结构关键点信息；

基于所述结构关键点信息和所述待测图像，确定所述待测图像中所述目标对象的分类。

2. 根据权利要求1所述的方法，其中所述目标对象包括虚拟人像，所述目标对象的结构关键点信息包括虚拟人体关键点信息。

3. 根据权利要求1所述的方法，其中基于所述结构关键点信息和所述待测图像，确定所述待测图像中所述目标对象的分类，包括：

基于所述待测图像生成第一目标特征，所述第一目标特征包括与所述目标对象关联的特征；

对所述结构关键点信息进行特征转换得到第二目标特征；

基于所述第一目标特征和所述第二目标特征生成第三目标特征；

基于所述第三目标特征确定所述待测图像中所述目标对象的分类。

4. 根据权利要求3所述的方法，其中所述基于所述第三目标特征确定所述待测图像中所述目标对象的分类，包括：

将所述第三目标特征输入分类网络，得到目标对象存在瑕疵置信度和/或目标对象不存在瑕疵置信度；

基于所述目标对象存在瑕疵置信度和/或所述目标对象不存在瑕疵置信度，得到目标对象的类别。

5. 根据权利要求3所述的方法，其中

所述方法还包括：在所述基于所述第三目标特征确定所述待测图像中所述目标对象的分类之前，对所述第三目标特征进行语义单元融合处理，以使所述第三目标特征中的语义单元能够融合所述第三目标特征中其他语义单元的信息；

所述基于所述第三目标特征确定所述待测图像中所述目标对象的分类，包括：基于语义单元融合处理后的第三目标特征，确定所述待测图像中所述目标对象的分类。

6. 根据权利要求1所述的方法，还包括：

响应于所述待测图像中不存在所述目标对象，确定所述待测图像不存在目标对象瑕

疵。

7. 一种图像检测装置，包括：

确定单元，用于确定待测图像中是否存在目标对象；

提取单元，用于响应于所述待测图像中存在所述目标对象，提取所述待测图像中所述目标对象的结构关键点信息；

分类单元，用于基于所述结构关键点信息和所述待测图像，确定所述待测图像中所述目标对象的分类。

8. 一种电子设备，包括：

至少一个存储器和至少一个处理器；

其中，所述存储器用于存储程序代码，所述处理器用于调用所述存储器所存储的程序代码以使所述电子设备执行权利要求 1 至 6 中任一项所述的方法。

9. 一种非暂态计算机存储介质，其中

所述非暂态计算机存储介质存储有程序代码，所述程序代码被计算机设备执行时，使得所述计算机设备执行权利要求 1 至 6 中任一项所述的方法。

10. 一种计算机程序产品，所述计算机程序产品包括程序代码，所述程序代码在被计算机设备执行时，使得所述计算机设备执行权利要求 1-6 中任一项所述的方法。