说 明 书

数据处理方法、装置、设备及可读存储介质

本申请要求于 2024 年 11 月 29 日提交中国专利局、申请号为 202411748954.0、申请名称为“数据处理方法、装置、设备及可读存储介质”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

本申请涉及计算机技术领域，尤其涉及一种数据处理方法、装置、设备及可读存储介质。

背景技术

在现有的图像问答领域中，用户上传的图像往往包含有隐私内容，当用户上传包含隐私内容的图像至模型进行图像问答时，模型在学习该图像和回答图像问题的过程中可能会导致隐私内容泄露，安全性低。

发明内容

本申请实施例提供了一种数据处理方法、装置、设备及可读存储介质，可以在保证图像问答的准确率同时，适当提高内容保护信息的安全性。

本申请实施例一方面提供了一种数据处理方法，包括：

基于图像以及图像对应的问题文本，在图像中确定问答注意力区域；问答注意力区域中的图像内容为图像中与问题文本相关联的图像内容；

获取内容保护信息，基于图像以及内容保护信息，在图像中确定内容保护区域；内容保护区域中的图像内容为图像中与内容保护信息相关联的图像内容；

若问答注意力区域与内容保护区域之间存在重叠区域，则基于问答注意力区域，在内容保护区域中确定第一待处理区域；第一待处理区域与问答注意力区域互不重叠；

其中，第一待处理区域用于指示对图像中的第一图像内容进行内容保护处理，第一图像内容为图像中第一待处理区域中的图像内容。

其中，基于问答注意力区域，在内容保护区域中确定第一待处理区域，包括：

将问答注意力区域与内容保护区域之间的重叠区域，确定为业务交叉区域，将内容保护区域中除业务交叉区域之外的区域，确定为第一待处理区域。

其中，还包括：

对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，在图像中通过第一处理内容替换掉第一图像内容，得到包括第一处理内容的问答保护图像；问答保护图像用于生成问题文本所对应的回答文本。

其中，还包括：

对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容；

在问答注意力区域中确定第二待处理区域，将第二待处理区域中的图像内容确定为第二图像内容，对第二图像内容进行内容保护处理，得到第二处理内容；

在图像中通过第一图像内容替换掉第一处理内容、且通过第二图像内容替换掉第二处理内容，得到包括第一处理内容和第二处理内容的问答保护图像。其中，在问答注意力区域中确定第二待处理区域，包括：

基于问题文本生成 T 个问题关键信息；T 为正整数，问题关键信息为用于识别得到问题文本的关键信息；

对问答注意力区域中的图像内容与 T 个问题关键信息进行注意力处理，得到问题注意力结果向量，基于问题注意力结果向量在问答注意力区域中确定与 T 个问题关键信息相关联的问题关键区域；

将问答注意力区域中除问题关键区域之外的区域，确定为第二待处理区域。

其中，基于图像以及图像对应的问题文本，在图像中确定问答注意力区域，包括：

对图像中的每个像素点进行连通域标记，得到 M 个连通域；M 为正整数，连通域用于表示图像中像素点的像素值满足连通域标记条件且像素点的位置相邻的图像区域；

对 M 个连通域分别进行特征提取，得到图像特征向量，对图像对应的问题文本进行特征提取，得到问题文本特征向量；

对图像特征向量与问题文本特征向量进行注意力处理，得到问答注意力结果向量，基于问答注意力结果向量在图像中确定问答注意力区域。

其中，基于问答注意力结果向量在图像中确定问答注意力区域，包括：

基于问答注意力结果向量生成 M 个连通域分别对应的注意力分值；

将具有最大注意力分值的连通域确定为目标连通域，基于目标连通域中像素点的位置信息生成边界轮廓点，将图像中边界轮廓点所包围的区域确定为问答注意力区域。

其中，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，包括：

获取内容保护区域相关联的保护程度分值；保护程度分值是基于内容保护信息对应的内容保护结果向量所生成的，内容保护结果向量是内容保护信息与图像进行注意力处理所得到的注意力向量；

基于保护程度分值确定目标重绘等级，在N个内容保护处理方式中确定目标重绘等级所指示的目标内容保护处理方式；N为正整数；

基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容。

其中，目标内容保护处理方式为图像隐身重绘；基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，包括：

将图像中除第一待处理区域之外的图像内容，确定为重绘源内容；

获取重绘源内容中的背景图像，对背景图像进行特征提取，得到背景特征向量；

基于背景特征向量，对第一待处理区域中的图像内容进行重绘处理，得到第一处理内容；第一处理内容的内容特征向量与背景特征向量具有相似关系。

其中，目标内容保护处理方式为图像模糊重绘；基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，包括：

获取模糊单元数值P，基于模糊单元数值P将第一待处理区域中的图像内容划分为P个待模糊图像单元；P为正整数；P个待模糊图像单元包括目标待模糊图像单元；

基于目标待模糊图像单元中每个像素点分别对应的颜色通道信息，生成目标待模糊图像单元的平均颜色通道信息，基于平均颜色通道信息对目标待模糊图像单元中的像素点进行模糊处理，得到已模糊图像单元；

当获取到P个待模糊图像单元分别对应的已模糊图像单元时，将P个已模糊图像单元确定为第一处理内容。

其中，目标内容保护处理方式为同类替换重绘；基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，包括：

基于内容保护信息生成K个对象关键信息；K为正整数，对象关键信息为用于识别得到内容保护信息的关键信息；

对第一待处理区域中的图像内容与K个对象关键信息进行注意力处理，得到对象注意力结果向量，基于对象注意力结果向量在第一待处理区域中确定与K个对象关键信息相关联的对象关键区域；

获取与内容保护信息相关联的内容提示文本，在内容提示文本中确定与对象关键区域相关联的目标提示文本，对目标提示文本进行特征提取，得到文本提示特征向量；

对对象关键区域对应的图像内容分别进行损坏处理，得到损坏内容，将图像中除第一待处理区域之外的图像内容，确定为重绘源内容，对重绘源内容进行特征提取，得到重绘源特征向量；

基于文本提示特征向量和重绘源特征向量对损坏内容进行图像修复，得到第一处理内容。

其中，目标内容保护处理方式为特征模糊重绘；基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，包括：

基于内容保护信息生成K个对象关键信息；K为正整数，对象关键信息为用于识别得到内容保护信息的关键信息；

对第一待处理区域中的图像内容与K个对象关键信息进行注意力处理，得到对象注意力结果向量，基于对象注意力结果向量在第一待处理区域中确定与K个对象关键信息相关联的对象关键区域，对对象关键区域对应的图像内容进行图像模糊重绘，得到第一处理内容。

本申请实施例一方面提供了一种数据处理装置，包括：

问答区域处理模块，用于基于图像以及图像对应的问题文本，在图像中确定问答注意力区域；问答注意力区域中的图像内容为图像中与问题文本相关联的图像内容；

保护区域处理模块，用于获取内容保护信息，基于图像以及内容保护信息，在图像中确定内容保护区域；内容保护区域中的图像内容为图像中与内容保护信息相关联的图像内容；

重绘区域处理模块，用于若问答注意力区域与内容保护区域之间存在重叠区域，则基于问答注意力区域，在内容保护区域中确定第一待处理区域；第一待处理区域与问答注意力区域互不重叠；其中，第一待处理区域用于指示对图像中的第一图像内容进行内容保护处理，第一图像内容为图像中第一待处理区域中

的图像内容。

在一种可能的实现方式中，重绘区域处理模块用于基于问答注意力区域，在内容保护区域中确定第一待处理区域时，具体用于执行以下操作：

将问答注意力区域与内容保护区域之间的重叠区域，确定为业务交叉区域，将内容保护区域中除业务交叉区域之外的区域，确定为第一待处理区域。

在一种可能的实现方式中，重绘区域处理模块还用于执行以下操作：

对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，在图像中通过第一处理内容替换掉第一图像内容，得到包括第一处理内容的问答保护图像；问答保护图像用于生成问题文本所对应的回答文本。

在一种可能的实现方式中，重绘区域处理模块还用于执行以下操作：

对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容；

在问答注意力区域中确定第二待处理区域，将第二待处理区域中的图像内容确定为第二图像内容，对第二图像内容进行内容保护处理，得到第二处理内容；

将所述第一图像内容已替换为所述第一处理内容、且所述第二图像内容已替换为所述第二处理内容的图像，确定为问答保护图像。

在一种可能的实现方式中，重绘区域处理模块用于在所述问答注意力区域中确定第二待处理区域时，具体用于执行以下操作：

基于问题文本生成 T 个问题关键信息；T 为正整数，问题关键信息为用于识别得到问题文本的关键信息；

对问答注意力区域中的图像内容与 T 个问题关键信息进行注意力处理，得到问题注意力结果向量，基于问题注意力结果向量在问答注意力区域中确定与 T 个问题关键信息相关联的问题关键区域；

将问答注意力区域中除问题关键区域之外的区域，确定为第二待处理区域。

在一种可能的实现方式中，问答区域处理模块用于基于图像以及图像对应的问题文本，在图像中确定问答注意力区域时，具体用于执行以下操作：

对图像中的每个像素点进行连通域标记，得到 M 个连通域；M 为正整数，连通域用于表示图像中像素点的像素值满足连通域标记条件且像素点的位置相邻的图像区域；

对 M 个连通域分别进行特征提取，得到图像特征向量，对图像对应的问题文本进行特征提取，得到问题文本特征向量；

对图像特征向量与问题文本特征向量进行注意力处理，得到问答注意力结果向量，基于问答注意力结果向量在图像中确定问答注意力区域。

在一种可能的实现方式中，问答区域处理模块用于基于问答注意力结果向量在图像中确定问答注意力区域时，具体用于执行以下操作：

基于问答注意力结果向量生成 M 个连通域分别对应的注意力分值；

将具有最大注意力分值的连通域确定为目标连通域，基于目标连通域中像素点的位置信息生成边界轮廓点，将图像中边界轮廓点所包围的区域确定为问答注意力区域。

在一种可能的实现方式中，重绘区域处理模块用于对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容时，具体用于执行以下操作：

获取内容保护区域相关联的保护程度分值；保护程度分值是基于内容保护信息对应的内容保护结果向量所生成的，内容保护结果向量是内容保护信息与图像进行注意力处理所得到的注意力向量；

基于保护程度分值确定目标重绘等级，在 N 个内容保护处理方式中确定目标重绘等级所指示的目标内容保护处理方式；N 为正整数；

基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容。

在一种可能的实现方式中，目标内容保护处理方式为图像隐身重绘；重绘区域处理模块用于基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容时，具体用于执行以下操作：

将图像中除第一待处理区域之外的图像内容，确定为重绘源内容；

获取重绘源内容中的背景图像，对背景图像进行特征提取，得到背景特征向量；

基于背景特征向量，对第一待处理区域中的图像内容进行重绘处理，得到第一处理内容；第一处理内容的内容特征向量与背景特征向量具有相似关系。

在一种可能的实现方式中，目标内容保护处理方式为图像模糊重绘；重绘区域处理模块用于基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容时，具体用于执行以下操作：

获取模糊单元数值P，基于模糊单元数值P将第一待处理区域中的图像内容划分为P个待模糊图像单元；P为正整数；P个待模糊图像单元包括目标待模糊图像单元；

基于目标待模糊图像单元中每个像素点分别对应的颜色通道信息，生成目标待模糊图像单元的平均颜色通道信息，基于平均颜色通道信息对目标待模糊图像单元中的像素点进行模糊处理，得到已模糊图像单元；

当获取到P个待模糊图像单元分别对应的已模糊图像单元时，将P个已模糊图像单元确定为第一处理内容。

在一种可能的实现方式中，目标内容保护处理方式为同类替换重绘；重绘区域处理模块用于基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容时，具体用于执行以下操作：

基于内容保护信息生成K个对象关键信息；K为正整数，对象关键信息为用于识别得到内容保护信息的关键信息；

对第一待处理区域中的图像内容与K个对象关键信息进行注意力处理，得到对象注意力结果向量，基于对象注意力结果向量在第一待处理区域中确定与K个对象关键信息相关联的对象关键区域；

获取与内容保护信息相关联的内容提示文本，在内容提示文本中确定与对象关键区域相关联的目标提示文本，对目标提示文本进行特征提取，得到文本提示特征向量；

对对象关键区域对应的图像内容分别进行损坏处理，得到损坏内容，将图像中除第一待处理区域之外的图像内容，确定为重绘源内容，对重绘源内容进行特征提取，得到重绘源特征向量；

基于文本提示特征向量和重绘源特征向量对损坏内容进行图像修复，得到第一处理内容。

在一种可能的实现方式中，目标内容保护处理方式为特征模糊重绘；重绘区域处理模块用于基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容时，具体用于执行以下操作：

基于内容保护信息生成K个对象关键信息；K为正整数，对象关键信息为用于识别得到内容保护信息的关键信息；

对第一待处理区域中的图像内容与K个对象关键信息进行注意力处理，得到对象注意力结果向量，基于对象注意力结果向量在第一待处理区域中确定与K个对象关键信息相关联的对象关键区域，对对象关键区域对应的图像内容进行图像模糊重绘，得到第一处理内容。

本申请实施例一方面提供了一种计算机设备，包括：处理器、存储器以及网络接口；

处理器与存储器、网络接口相连，其中，网络接口用于提供数据通信功能，存储器用于存储计算机程序，计算机程序被处理器执行时，使得该计算机设备执行本申请实施例提供的方法。

本申请实施例一方面提供了一种计算机可读存储介质，计算机可读存储介质存储有计算机程序，该计算机程序适于由处理器加载并执行，以使得具有该处理器的计算机设备执行本申请实施例提供的方法。

本申请实施例一方面提供了一种计算机程序产品，该计算机程序产品包括计算机程序，该计算机程序存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机程序，处理器执行该计算机程序，使得该计算机设备执行本申请实施例提供的方法。

本申请实施例通过确定问答注意力区域和内容保护区域，可以区分模型处理图像问答所需的图像内容，进而基于问答注意力区域，在内容保护区域中确定第一待处理区域，对第一待处理区域中的第一图像内容进行内容保护处理，减少了模型在学习图像和回答问题文本的过程中，可能会导致与内容保护信息相关联的图像内容泄露的情况发生，同时，由于第一待处理区域与问答注意力区域互不重叠，所以模型依旧可以通过问答注意力区域识别到与问题文本相关联的图像内容，可见本申请可以在保证图像问答的准确率同时，适当提高内容保护信息的安全性。

附图说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图1是本申请实施例提供的一种网络架构示意图；

图 2 是本申请实施例提供的一种数据处理场景示意图一；
图 3 是本申请实施例提供的一种数据处理方法的流程示意图一；
图 4 是本申请实施例提供的一种数据处理方法的流程示意图二；
图 5 是本申请实施例提供的一种数据处理场景示意图二；
图 6 是本申请实施例提供的一种数据处理场景示意图三；
图 7 是本申请实施例提供的一种数据处理装置的结构示意图；
图8是本申请实施例提供的一种计算机设备的结构示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

请参见图 1，图 1 是本申请实施例提供的一种网络架构示意图。如图 1 所示，该网络架构可以包括终端设备 100 和业务服务器 200，终端设备 100 可以与业务服务器 200 之间存在通信连接，上述通信连接不限定连接方式，可以通过有线通信方式进行直接或间接地连接，也可以通过无线通信方式进行直接或间接地连接，还可以通过其它方式，本申请在此不做限制。

其中，终端设备可以包括：智能手机、平板电脑、笔记本电脑、台式计算机、智能语音交互设备、智能家电（例如，智能电视）、可穿戴设备、车载终端、飞行器等具有数据处理功能的智能终端。其中，车载终端可以为智慧交通场景和辅助驾驶场景下的终端设备。应当理解，如图 1 所示的终端设备 100 可以安装有具备数据处理功能的应用客户端，当应用客户端运行于各终端设备中时，可以分别与上述图 1 所示的业务服务器 200 之间进行数据交互。

其中，应用客户端具体可以包括：车载客户端、智能家居客户端、娱乐客户端（例如，游戏客户端）、多媒体客户端（例如，视频客户端）、社交客户端以及资讯类客户端（例如，新闻客户端）等。其中，本申请实施例中的应用客户端可以集成在某客户端（例如，社交客户端）中，也可以为独立的客户端（例如，新闻客户端），本申请实施例不对应用客户端的类型进行限定。

其中，业务服务器 200 可以为应用客户端对应的服务器，业务服务器 200 可以是独立的物理服务器，也可以是多个物理服务器构成的服务器集群或者分布式系统，还可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN（Content Delivery Network，内容分发网络）、以及大数据和人工智能平台等基础云计算服务的云服务器。

如图 1 所示，对象（用户）可以在终端设备 100 中通过应用客户端上传图像、图像对应的问题文本以及内容保护信息，终端设备 100 中可以部署有端侧多模态模型。其中，端侧多模态模型可以是指用于在终端设备上运行的，能够同时处理多种模态数据（如文本、图像、音频等）的人工智能模型。内容保护信息可以是针对图像中的隐私的图像内容的描述信息，内容保护信息可以是文本数据、图像数据或者音频数据。

终端设备 100 可以通过端侧多模态模型对图像与问题文本进行注意力处理（即可以学习图像与问题文本之间的相关性），以在图像中确定与问题文本相关联的问答注意力区域。同样的，终端设备 100 可以通过端侧多模态模型对图像与内容保护信息进行注意力处理（即可以学习图像与内容保护信息之间的相关性），以在图像中确定与内容保护信息相关联的内容保护区域。问答注意力区域和内容保护区域均可以是一个通过多边形顶点描述的图像遮罩。

若问答注意力区域与内容保护区域之间存在重叠区域，则终端设备 100 可以将问答注意力区域与内容保护区域之间的重叠区域，确定为业务交叉区域，将内容保护区域中除业务交叉区域之外的区域，确定为第一待处理区域。通过端侧多模态模型对第一待处理区域中的第一图像内容进行内容保护处理，得到第一处理内容。内容保护处理包括内容隐藏处理、重绘处理等消除内容保护信息所指示的图像等方式，内容隐藏处理可以是指添加不透明的遮罩层到图像区域，从而隐藏该图像区域。也可以是直接擦除（消除）该图像区域。重绘处理可以包括图像隐身重绘、图像模糊重绘、同类替换重绘和特征模糊重绘等，本申请实施例在此不做限制。

将第一图像内容已替换为第一处理内容的图像，确定为问答保护图像。其中，第一图像内容为图像中第一待处理区域中未被重绘的图像内容。

终端设备 100 可以将经过内容保护处理得到的问答保护图像以及问题文本发送至业务服务器 200，业务服务器通过云端多模态模型对问答保护图像进行识别，生成问题文本对应的回答文本，将回答文本返回至终端设备 100。

可以理解，终端设备100中的端侧多模态模型也可以部署在业务服务器200中，业务服务器200可以获取图像，通过端侧多模态模型对图像进行内容保护处理，得到问答保护图像，再通过云端多模态模型对问答保护图像进行识别，生成问题文本对应的回答文本。端侧多模态模型和云端多模态模型可以通过模型隔离以保证内容保护信息的安全性，上述模型隔离可以是通过物理隔离或者可信环境隔离（可信环境需授权才可访问），本申请实施例在此不做限制。

本申请实施例通过确定问答注意力区域和内容保护区域，可以区分模型处理图像问答所需的图像内容，而基于问答注意力区域，在内容保护区域中确定第一待处理区域，通过部署在终端设备100的端侧多模态模型对第一待处理区域中的第一图像内容进行内容保护处理，使得内容保护信息所指示的图像内容无法被模型识别，进而将问答保护图像以及问题文本发送至业务服务器200的云端多模态模型时，减少了模型在学习图像和回答问题文本的过程中，可能会导致与内容保护信息相关联的图像内容泄露的情况发生，同时，由于第一待处理区域与问答注意力区域互不重叠，所以模型依旧可以通过问答注意力区域识别到与问题文本相关联的图像内容，可见本申请可以在保证图像问答的准确率同时，适当提高内容保护信息的安全性。

请参见图2，图2是本申请实施例提供的一种数据处理的场景示意图一。如图2所示，对象（用户）可以在终端设备中通过应用客户端上传图像A以及图像A对应的问题文本。其中，图像A对应的问题文本可以是通过音频识别对象输入的语音得到的文本或者是对象在应用客户端直接输入的文本，本申请实施例在此不做限制。终端设备可以是上述图1对应实施例的终端设备100。问题文本例如可以是“图像中的汽车的类型”。

终端设备可以获取内容保护信息，内容保护信息可以是针对图像A中的隐私的图像内容的描述信息，内容保护信息可以是一种类型描述或者直接指定的特定描述，可以是在应用客户端中的隐私内容类型里进行选择配置，也可以由对象输入，也可以是模型自动识别得到的。内容保护信息可以是文本数据、图像数据或者音频数据。内容保护信息的类型可以包括个人身份信息、联系信息、财务信息、车辆信息、医疗信息和地址位置信息等。内容保护信息例如可以是“人物/人像”。

终端设备中可以部署有端侧多模态模型。其中，端侧多模态模型可以是指用于在终端设备上运行的，能够同时处理多种模态数据（如文本、图像、音频等）的人工智能模型，例如可以是Phi-3-mini-128k-instruct模型（一种开源轻量级模型），Chat-GLM（General Language Model）模型（一种自然语言生成模型）、Qwen-VLM（Vision Language Model）模型（一种大规模视觉语言模型）等可以在边缘设备上运行的多模态大语言模型。

终端设备可以通过端侧多模态模型、图像A与问题文本，在图像A中确定与问题文本相关联的问答注意力区域B。同样的，终端设备可以通过端侧多模态模型、图像A与内容保护信息，在图像A中确定与内容保护信息相关联的内容保护区域C。问答注意力区域B和内容保护区域C均可以是一个通过多边形顶点描述的图像遮罩。

如图2所示，问答注意力区域B与内容保护区域C之间存在重叠区域，终端设备可以将问答注意力区域与内容保护区域之间的重叠区域，确定为业务交叉区域D，业务交叉区域D可以是斜线表示的区域。终端设备可以将内容保护区域C中除业务交叉区域D之外的区域，确定为第一待处理区域E，第一待处理区域E可以是横线表示的区域。

终端设备可以对第一待处理区域E进行内容保护处理，内容保护处理包括内容隐藏处理、重绘处理等消除内容保护信息所指示的图像等方式，内容隐藏处理可以是指添加不透明的遮罩层到图像区域，从而隐藏该图像区域。也可以是直接擦除（消除）该图像区域。重绘处理可以包括图像隐身重绘、图像模糊重绘、同类替换重绘和特征模糊重绘等，本申请实施例在此不做限制。

以内容保护处理为重绘处理为例，终端设备可以通过内容保护区域C中的图像内容与内容保护信息的关联程度确定保护程度分值，通过保护程度分值确定内容保护处理的内容保护处理方式。

以内容保护处理为图像隐身重绘为例，图像隐身重绘可以是基于图像A中的背景图像（例如可以是树木）在第一待处理区域E中生成第一处理内容F，将第一图像内容已替换为第一处理内容F的图像A，确定为问答保护图像G。其中，第一图像内容为图像中第一待处理区域E中的图像内容，问答保护图像G用于进行图像识别生成问题文本对应的回答文本。问答保护图像G可以通过第一处理内容F对内容保护信息相关联的图像内容起到安全保护作用。

本申请实施例可以应用于图像问答任务（Visual Question Answering，VQA）和图像重绘任务（Visual Re-drawing），在图像重绘任务中，本申请可以通过内容保护信息，对所指示的图像内容进行局部重绘，实现部分图像内容的风格转换，同时保留图像中其他的原始特征，可以实现对图像的多样化处理，可以应用

于图像编辑、视觉设计等领域。在图像问答任务中，本申请可以对第一待处理区域中的第一图像内容进行内容保护处理，减少了模型在学习图像和回答问题文本的过程中，可能会导致与内容保护信息相关联的图像内容泄露的情况发生，从而有效地保护内容保护信息，同时保留图像中的与问题文本相关联的图像内容，保证图像问答的准确率不受影响。

请参见图 3，图 3 是本申请实施例提供的一种数据处理方法的流程示意图一，该数据处理方法可以由计算机设备执行，计算机设备可以是如图 1 所示的终端设备 100 或业务服务器 200。以下将以本数据处理方法由计算机设备执行为例进行说明。其中，该数据处理方法至少可以包括以下步骤 S101-步骤 S103：

步骤 S101，基于图像以及图像对应的问题文本，在图像中确定问答注意力区域；问答注意力区域中的图像内容为图像中与问题文本相关联的图像内容；

具体的，计算机设备可以获取图像以及图像对应的问题文本，图像对应的问题文本可以是通过音频识别对象输入的语音得到的文本或者是对象在应用客户端直接输入的文本，本申请实施例在此不做限制。问题文本例如可以是“图像中车辆的类型”、“图像是否包含车辆”、“图像拍摄的时间”等。

计算机设备可以部署有端侧多模态模型。其中，端侧多模态模型可以是指用于在终端设备上运行的，能够同时处理多种模态数据（如文本、图像、音频等）的人工智能模型，例如可以是 Phi-3-mini-128k-instruct 模型，Chat-GLM 模型、Qwen-VLM 模型等可以在边缘设备上运行的多模态大语言模型。

计算机设备可以通过端侧多模态模型将图像划分为多个图像区域，例如为 M 个图像区域，图像区域可以是连通域、形状域。对多个图像区域与问题文本进行注意力处理，得到问答注意力结果向量，问答注意力结果向量可以是一个 M 维向量，问答注意力结果向量中的每一维可以分别代表第一图像区域与问题文本的第一注意力分值，第一注意力分值可以表示该图像区域与问题文本的相关程度，计算机设备可以将具有最大第一注意力分值的图像区域确定为与问题文本相关联的问答注意力区域。其中，问答注意力区域可以是一个通过多边形顶点描述的图像遮罩。可选的，计算机设备可以将多个第一注意力分值达到第一分值阈值（预先设置的数值）的图像区域分别确定为问答注意力区域，即可以获得多个问答注意力区域。

步骤 S102，获取内容保护信息，基于图像以及内容保护信息，在图像中确定内容保护区域；内容保护区域中的图像内容为图像中与内容保护信息相关联的图像内容；

具体的，计算机设备可以获取内容保护信息。内容保护信息可以是针对图像中的隐私的图像内容的描述信息，内容保护信息可以是一种类型描述或者直接指定的特定描述，可以是在应用客户端中的隐私内容类型里进行选择配置，也可以由对象输入，也可以是模型自动识别得到的。内容保护信息可以是文本数据、图像数据或者音频数据。内容保护信息的类型可以包括个人身份信息、联系信息、财务信息、车辆信息、医疗信息和地址位置信息等。

计算机设备可以通过端侧多模态模型对多个图像区域与内容保护信息进行注意力处理，得到内容保护结果向量，内容保护结果向量也可以是一个 M 维向量，内容注意力结果向量中的每一维可以分别代表一个图像区域与内容保护信息的第二注意力分值，第二注意力分值可以表示该图像区域与内容保护信息的相关程度，计算机设备可以将具有最大第二注意力分值的图像区域确定为与内容保护信息相关联的内容保护区域。其中，内容保护区域可以是一个通过多边形顶点描述的图像遮罩。可选的，计算机设备可以将多个第二注意力分值达到第二分值阈值（预先设置的数值）的图像区域分别确定为内容保护区域，即可以获得多个内容保护区域。

步骤 S103，若问答注意力区域与内容保护区域之间存在重叠区域，则基于问答注意力区域，在内容保护区域中确定第一待处理区域；第一待处理区域与问答注意力区域互不重叠；其中，第一待处理区域用于指示对图像中的第一图像内容进行内容保护处理，第一图像内容为图像中第一待处理区域中的图像内容。

具体的，若问答注意力区域与内容保护区域之间存在重叠区域，则计算机设备可以将问答注意力区域与内容保护区域之间的重叠区域，确定为业务交叉区域，将内容保护区域中除业务交叉区域之外的区域，确定为第一待处理区域。

计算机设备可以对第一待处理区域进行内容保护处理，内容保护处理包括内容隐藏处理、重绘处理等消除内容保护信息所指示的图像等方式，内容隐藏处理可以是指添加不透明的遮罩层到图像区域，从而隐藏该图像区域。也可以是直接擦除（消除）该图像区域。重绘处理可以包括图像隐身重绘、图像模糊重绘、同类替换重绘和特征模糊重绘等，本申请实施例在此不做限制。

计算机设备可以通过内容保护区域对应的第二注意力分值确定保护程度分值，例如可以将第二注意力分值确定为内容保护区域的保护程度分值，通过保护程度分值确定内容保护处理的内容保护处理方式。

计算机设备可以通过端侧多模态模型对第一待处理区域对应的第一图像内容进行内容保护处理，得到

第一处理内容，将第一图像内容已替换为第一处理内容的图像，确定为问答保护图像。其中，第一图像内容为图像中第一待处理区域中的图像内容，问答保护图像可以用于生成问题文本对应的回答文本。

可以理解，计算机设备可以在图像中确定一个或多个第一待处理区域，每个第一待处理区域可以对应有不同的保护程度分值，即每个第一待处理区域的重绘方式可以互不相同。

本申请实施例通过确定问答注意力区域和内容保护区域，可以区分模型处理图像问答所需的图像内容，进而基于问答注意力区域，在内容保护区域中确定第一待处理区域，对第一待处理区域中的第一图像内容进行内容保护处理，减少了模型在学习图像和回答问题文本的过程中，可能会导致与内容保护信息相关联的图像内容泄露的情况发生，同时，由于第一待处理区域与问答注意力区域互不重叠，所以模型依旧可以通过问答注意力区域识别到与问题文本相关联的图像内容，可见本申请可以在保证图像问答的准确率同时，适当提高内容保护信息的安全性。

请参见图 4，图 4 是本申请实施例提供的一种数据处理方法的流程示意图二，该数据处理方法可以由计算机设备执行，计算机设备可以是如图 1 所示的终端设备 100 或业务服务器 200。以下将以本数据处理方法由计算机设备执行为例进行说明。其中，该数据处理方法至少可以包括以下步骤 S201-步骤 S206：

步骤 S201，对图像中的每个像素点进行连通域标记，得到 M 个连通域；M 为正整数，连通域用于表示图像中像素点的像素值满足连通域标记条件且像素点的位置相邻的图像区域；对 M 个连通域分别进行特征提取，得到图像特征向量，对图像对应的问题文本进行特征提取，得到问题文本特征向量；对图像特征向量与问题文本特征向量进行注意力处理，得到问答注意力结果向量；

具体的，计算机设备可以部署有端侧多模态模型。其中，端侧多模态模型可以是指用于在终端设备上运行的，能够同时处理多种模态数据（如文本、图像、音频等）的人工智能模型，例如可以是 Phi-3-mini-128k-instruct 模型，Chat-GLM 模型、Qwen-VLM 模型等可以在边缘设备上运行的多模态大语言模型。端侧多模态模型可以是一种图文多模态模型，图文多模态模型是一种能够同时处理和理解图像和文本数据的人工智能模型。这类模型通过结合视觉和语言信息，能够执行诸如图像描述生成、图像问答、图像分类与文本匹配等任务。其核心在于融合和对齐不同模态的数据，使得模型能够从多种信息源中提取有用的特征，并进行综合分析和推理。

计算机设备可以将图像输入至端侧多模态模型的图像处理层，端侧多模态模型的图像处理层可以是 CLIP（Contrastive Language-Image Pre-Training）中的图像编码器或者 ViT（Vision Transformer）图像编码器，本申请实施例在此不做限制。计算机设备可以通过图像处理层对图像进行二值化处理，将图像中大于或等于业务灰度阈值的像素点的灰度值设置为 255，将图像中小于业务灰度阈值的像素点的灰度值设置为 0，得到图像的灰度图。其中，业务灰度阈值可以通过业务需求或者具体的图像进行设置，本申请实施例在此不做限制。

计算机设备可以对图像的灰度图进行连通域标记，将满足连通域标记条件并且像素点的位置相邻的像素点标记为一个连通域，得到 M 个连通域。其中，连通域标记条件可以是指像素点的灰度值相同。连通域用于表示图像中像素点的像素值满足连通域标记条件且像素点的位置相邻的图像区域。计算机设备可以对 M 个连通域中的图像内容分别进行特征提取，得到图像特征向量。

计算机设备可以将问题文本输入至端侧多模态模型的文本处理层，端侧多模态模型的文本处理层可以是 BERT（Bidirectional Encoder Representations from Transformers）文本编码器，计算机设备可以通过文本处理层中的 Tokenizer（分词器）将问题文本拆分编码为词元（token），例如问题文本为“图像中的汽车类型”，该问题文本对应的词元序列可以表示为[图像,汽车,类型]。其中，[图像]、[汽车和[类型]均为词元。文本拆分的方法可以是基于单词拆分（word-based）、基于字符拆分（character-based）、基于子词拆分（subword-based），本申请实施例在此不做限制。计算机设备可以将问题文本中的所有 token 可以被嵌入（embedding）成相同大小的向量，然后将这些相同大小的向量进行向量拼接，得到问题文本特征向量。

计算机设备可以将图像特征向量与问题文本特征向量输入至端侧多模态模型的注意力处理层，将问题文本特征向量与查询参数矩阵W_i^Q进行点乘运算，得到查询向量 Q，将图像特征向量与键参数矩阵W_i^K进行点乘运算，得到键向量 K，将图像特征向量与与值参数W_i^V矩阵进行点乘运算，得到值向量 V。其中，查询参数矩阵W_i^Q、键参数矩阵W_i^K与值参数矩阵W_i^V均是由注意力处理的可学习参数所构成的矩阵。

计算机设备可以将查询向量Q与键向量 K 的转置运算结果K^T进行点乘运算，得到注意力分数向量。基于键向量 K 的维度值d_K的平方根$\sqrt{d_K}$对注意力分数向量进行降维处理，对降维后的注意力分数向量进行归

一化处理，得到注意力权重向量，将注意力权重向量与值向量进行点乘运算得到问答注意力结果向量$Attn_Q$，其过程可以如公式（1）所示：

$$Attn_Q = \text{softmax}\left(\frac{Q \cdot K^T}{\sqrt{d_k}}\right) \cdot V \quad \text{公式（1）}$$

其中，softmax函数为归一化处理，d_K为键向量 K 的维度值。

步骤 S202，基于问答注意力结果向量生成 M 个连通域分别对应的注意力分值；将具有最大注意力分值的连通域确定为目标连通域，基于目标连通域中像素点的位置信息生成边界轮廓点，将图像中边界轮廓点所包围的区域确定为问答注意力区域。

具体的，问答注意力结果向量$Attn_Q$可以是一个 M 维向量，问答注意力结果向量$Attn_Q$可以中的每一维可以分别代表一个连通域与问题文本的第一注意力分值，第一注意力分值可以表示该连通域与问题文本的相关程度，计算机设备可以将具有最大第一注意力分值的连通域确定为目标连通域，计算机设备可以基于目标连通域中像素点的位置信息，确定 S 个边界轮廓点。边界轮廓点的数量可以是基于目标连通域的范围以及形状所确定的。计算机设备可以将边界轮廓点所组成的多边形区域确定为问答注意力区域。其中，问答注意力区域可以是一个通过多边形顶点描述的图像遮罩。例如问答注意力区域可以以文本形式表示多边形顶点坐标，例如为{[100, 50], [50, 200], [150, 200]}，其中，[100, 50], [50, 200]和[150, 200]均为边界轮廓点在图像中的坐标。可选的，计算机设备可以将多个第一注意力分值达到第一分值阈值（预先设置的数值）的图像区域分别确定为问答注意力区域，即可以获得多个问答注意力区域。

步骤 S203，获取内容保护信息，基于图像以及内容保护信息，在图像中确定内容保护区域；内容保护区域中的图像内容为图像中与内容保护信息相关联的图像内容；

具体的，计算机设备可以获取内容保护信息。内容保护信息可以是针对图像中的隐私的图像内容的描述信息，内容保护信息可以是预设的一种类型描述或者直接指定的特定描述，可以是在应用客户端中的隐私内容类型里进行选择配置，也可以由对象输入，也可以是模型自动识别得到的。内容保护信息可以是文本数据、图像数据或者音频数据。内容保护信息的类型可以包括个人身份信息、联系信息、财务信息、车辆信息、医疗信息和地址位置信息等。

计算机设备可以对图像数据进行图像识别或者对音频数据进行音频识别，得到内容保护信息的文本数据，以内容保护信息为文本数据为例，计算机设备可以通过端侧多模态模型的文本处理层对内容保护信息进行特征提取，得到内容保护信息对应的内容保护特征向量，其过程可以参见上述步骤 S201 中文本处理层的具体描述，在此不再赘述。

计算机设备可以将图像特征向量与内容保护特征向量输入至端侧多模态模型的注意力处理层，通过注意力处理层对图像特征向量与内容保护特征向量进行交叉注意力计算，得到内容保护结果向量。交叉注意力计算的过程可以参见上述步骤 S201 中注意力处理层的具体描述，在此不再赘述。

内容保护结果向量也可以是一个 M 维向量，内容注意力结果向量中的每一维可以分别代表一个图像区域与内容保护信息的第二注意力分值，第二注意力分值可以表示该图像区域与内容保护信息的相关程度，计算机设备可以通过内容保护结果向量中的第二注意力分值，在 M 个连通域中确定与问题文本相关联的业务连通域，基于业务连通域在图像中，确定内容保护区域。其中，内容保护区域中的图像内容为图像中与内容保护信息相关联的图像内容。确定业务连通域以及内容保护区域的过程可以参见步骤 S202 中关于目标连通域以及问答注意力区域的具体描述，在此不再赘述。可选的，计算机设备可以将多个第二注意力分值达到第二分值阈值（预先设置的数值）的图像区域分别确定为内容保护区域，即可以获得多个内容保护区域。

步骤 S204，将问答注意力区域与内容保护区域之间的重叠区域，确定为业务交叉区域，将内容保护区域中除业务交叉区域之外的区域，确定为第一待处理区域。

具体的，计算机设备可以通过问答注意力区域的边界轮廓点与内容保护区域的边界轮廓点，判断问答注意力区域与内容保护区域之间是否重叠，若问答注意力区域与内容保护区域之间存在重叠区域，则计算机设备可以基于问答注意力区域的边界轮廓点与内容保护区域的边界轮廓点确定重叠轮廓点，将重叠轮廓点所组成的多边形区域确定为重叠区域。计算机设备可以将问答注意力区域与内容保护区域之间的重叠区

域，确定为业务交叉区域，将内容保护区域中除业务交叉区域之外的区域，确定为第一待处理区域。

步骤S205，获取内容保护区域相关联的保护程度分值；保护程度分值是基于内容保护信息对应的内容保护结果向量所生成的，内容保护结果向量是内容保护信息与图像进行注意力处理所得到的注意力向量；基于保护程度分值确定目标重绘等级，在N个内容保护处理方式中确定目标重绘等级所指示的目标内容保护处理方式；N为正整数；基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容。

具体的，计算机设备可以对第一待处理区域进行内容保护处理，内容保护处理包括内容隐藏处理、重绘处理等消除内容保护信息所指示的图像等方式，内容隐藏处理可以是指添加不透明的遮罩层到图像区域，从而隐藏该图像区域。也可以是直接擦除（消除）该图像区域。重绘处理可以包括图像隐身重绘、图像模糊重绘、同类替换重绘和特征模糊重绘等，本申请实施例在此不做限制。

计算机设备可以将内容保护结果向量中的M个注意力分值进行归一化处理，得到每个注意力分值对应的归一化分值。计算机设备可以业务连通域在内容保护结果向量中对应的注意力分值对应的归一化分值，确定为内容保护区域相关联的保护程度分值。保护程度分值可以为0到1的分值，例如可以为0.7。计算机设备可以通过将保护程度分值设置为内容保护区域的图像遮罩的灰度值来进行记录。计算机设备在得到保护程度分值后，该保护程度分值可以作为灰度值存储在作为mask的内容保护区域中，在下文确定内容保护处理方式时，计算机设备可以通过提取mask中的灰度值确定保护程度分值。

计算机设备可以基于保护程度分值确定内容保护区域的目标重绘等级，在N个内容保护处理方式中确定目标重绘等级所指示的目标内容保护处理方式。以内容保护处理为重绘处理为例，保护程度分值与内容保护处理方式的映射关系与如表1所示：

表1

重绘制等级	保护程度分值	内容保护处理方式
T0	[0,0.2)	维持原图
T1	[0.2,0.4)	特征模糊重绘
T2	[0.4,0.6)	同类替换重绘
T3	[0.6,0.8)	图像模糊重绘
T4	[0.8,1]	图像隐身重绘

例如，若保护程度分值为0.1，则计算机设备可以确定重绘制等级为T0，对对第一待处理区域中的图像内容不进行重绘，维持原图。若保护程度分值为0.3，则计算机设备可以确定重绘制等级为T1，对对第一待处理区域中的图像内容进行特征模糊重绘。若保护程度分值为0.5，则计算机设备可以确定重绘制等级为T2，对对第一待处理区域中的图像内容进行同类替换重绘。若保护程度分值为0.7，则计算机设备可以确定重绘制等级为T3，对对第一待处理区域中的图像内容进行图像模糊重绘。若保护程度分值为0.9，则计算机设备可以确定重绘制等级为T4，对对第一待处理区域中的图像内容进行图像隐身重绘。其中，第一图像内容为图像中第一待处理区域中的图像内容。

再请一并参加图5，图5是本申请实施例提供的一种数据处理场景示意图二，如图5所示，以图像A0，第一待处理区域中的图像内容为第一图像内容P0为例，计算机设备可以基于目标内容保护处理方式，对第一图像内容P0进行内容保护处理。

当目标内容保护处理方式为图像隐身重绘时，基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容的过程可以是：将图像中除第一待处理区域之外的图像内容，确定为重绘源内容；获取重绘源内容中的背景图像，对背景图像进行特征提取，得到背景特征向量；基于背景特征向量，对第一待处理区域中的图像内容进行重绘处理，得到第一处理内容；第一处理内容的内容特征向量与背景特征向量具有相似关系。

具体的，计算机设备可以将图像A0中除第一待处理区域之外的图像内容，确定为重绘源内容，计算机设备可以对重绘源内容进行背景识别，得到重绘源内容中的背景图像。计算机设备可以将背景图像输入至去噪扩散概率模型（Denoising Diffusion Probabilistic Models，DDPM）或者潜在扩散模型（Latent Stable Diffusion Models，LDM），通过ControlNet（一种将条件控制添加到图像扩散模型中的神经网络）和IPAdapter（Text Compatible Image Prompt Adapter for Text-to-Image Diffusion Models，一种用于实现图像提示的适配器）将背景图像对应的背景特征向量确定为噪声图在后向扩散过程的条件输入，从而生成与背景图像具有相似关系的第一处理内容P4。其中，第一处理内容P4的内容特征向量与背景特征

向量具有相似关系。上述过程也可以称之为通过图像扩散模型对第一待处理区域中的图像内容进行OutPainting（图像外部填充）。计算机设备可以将第一图像内容P0已替换为第一处理内容P4的图像，确定为问答保护图像A4。

可选的，图像隐身重绘可以是直接将第一图像内容P0的像素点的像素值设置为0或255，使得第一图像内容P0不可见，从而直接进行图像隐身。也可以通过背景替换（Background Removal）的方式对第一图像内容P0进行图像内容的替换，背景替换是一种通过替换图像中的背景来隐藏或去除敏感信息的技术。这种技术可以有效地保护隐私，同时保留图像中的主要对象或前景信息。

当目标内容保护处理方式为图像模糊重绘时，基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容的过程可以是：获取模糊单元数值P，基于模糊单元数值P将第一待处理区域中的图像内容划分为P个待模糊图像单元；P为正整数；P个待模糊图像单元包括目标待模糊图像单元；基于目标待模糊图像单元中每个像素点分别对应的颜色通道信息，生成目标待模糊图像单元的平均颜色通道信息，基于平均颜色通道信息对目标待模糊图像单元中的像素点进行模糊处理，得到已模糊图像单元；当获取到P个待模糊图像单元分别对应的已模糊图像单元时，将P个已模糊图像单元确定为第一处理内容。

具体的，计算机设备可以获取模糊单元数值P，模糊单元数值P可以是基于第一图像内容P0的范围大小所确定的，第一图像内容P0的范围越大，模糊单元数值P越大。计算机设备可以基于模糊单元数值P将第一图像内容P0划分为P个待模糊图像单元，P个待模糊图像单元包括目标待模糊图像单元。

图像模糊重绘的方式可以包括高斯模糊(Gaussian Blur)、均值模糊(Mean Blur)、双边滤波(Bilateral Filtering）等。其中，图像模糊化脱敏技术（Blurring）：是一种常用的图像处理技术，主要用于保护隐私和敏感信息。通过模糊化处理，可以使图像中的某些部分变得不清晰，从而防止未经授权的人员获取图像中的详细信息。

以均值模糊为例，计算机设备可以基于目标待模糊图像单元中每个像素点分别对应的颜色通道信息，生成目标待模糊图像单元的平均颜色通道信息，基于平均颜色通道信息对目标待模糊图像单元中的像素点进行模糊处理，例如可以是将平均颜色通道信息应用至目标待模糊图像单元中的每个像素点，得到已模糊图像单元。当获取到P个待模糊图像单元分别对应的已模糊图像单元时，计算机设备可以将P个已模糊图像单元确定为第一处理内容P3。计算机设备可以将第一图像内容P0已替换为第一处理内容P3的图像，确定为问答保护图像A3。

可选的，图像模糊重绘的方式还可以是图像卡通化（Cartooning）或色块填充（Color Filling），其中，图像卡通化是一种将图像转换为卡通风格的图像处理技术。通过这种技术，可以将图像中的细节和纹理简化，使其看起来像手绘的卡通图像。这种技术不仅可以用于艺术创作，还可以用于隐私保护和数据脱敏，因为卡通化处理可以模糊或简化图像中的敏感信息，使其难以识别。色块填充是一种通过用随机颜色块或噪声填充图像中某些区域来隐藏敏感信息的技术。这种方法不仅改变了颜色，还引入了随机性，使得原始信息更难以恢复，从而有效地保护隐私。

当目标内容保护处理方式为同类替换重绘时，基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容的过程可以是：基于内容保护信息生成K个对象关键信息；K为正整数，对象关键信息为用于识别得到内容保护信息的关键信息；对第一待处理区域中的图像内容与K个对象关键信息进行注意力处理，得到对象注意力结果向量，基于对象注意力结果向量在第一待处理区域中确定与K个对象关键信息相关联的对象关键区域；获取与内容保护信息相关联的内容提示文本，在内容提示文本中确定与对象关键区域相关联的目标提示文本，对目标提示文本进行特征提取，得到文本提示特征向量；对对象关键区域对应的图像内容分别进行损坏处理，得到损坏内容，将图像中除第一待处理区域之外的图像内容，确定为重绘源内容，对重绘源内容进行特征提取，得到重绘源特征向量；基于文本提示特征向量和重绘源特征向量对损坏内容进行图像修复，得到第一处理内容。

具体的，计算机设备可以基于内容保护信息生成K个对象关键信息，对象关键信息可以是用于识别得到内容保护信息的关键信息。例如当内容保护信息为“女性人物”时，对象关键信息可以为“眼睛（虹膜）”、“鼻子”、“嘴唇（唇纹）”等。计算机设备可以对第一图像内容P0与K个对象关键信息进行注意力处理，得到对象注意力结果向量，基于对象注意力结果向量在第一待处理区域中确定与K个对象关键信息相关联的对象关键区域，例如可以得到对象关键信息在第一图像内容P0的图像区域，对象关键区域中的图像内容可以用于表示对象关键信息。例如对象关键区域可以是第一图像内容P0中表示人物的眼睛、人物的鼻子和人物的嘴巴的图像区域。

计算机设备可以获取与对象关键区域相关联的若干个内容提示文本(prompt),例如对象关键区域为人物的嘴巴的图像区域（对象关键信息为嘴巴）时，计算机设备可以获取与人物的嘴巴相关联的提示词，例如可以包括“微笑的嘴”、“张开的嘴”等。计算机设备可以对若干个内容提示文本与对象关键区域所表示的对象关键信息进行注意力处理，得到对象关键结果向量，基于对象关键结果向量在若干个内容提示文本中确定与对象关键区域相关联的目标提示文本。例如目标提示文本可以是“微笑的嘴”，目标提示文本也可以与内容保护信息的文本相同。可选的，计算机设备也可以通过内容保护区域的图像内容进行图像识别，生成目标提示文本。

计算机设备可以对对象关键区域对应的图像内容(例如可以是人物的嘴巴的图像区域)进行损坏处理，得到损坏内容。损坏处理可以是对对象关键区域对应的图像内容添加符合高斯分布的噪声。

计算机设备可以将目标提示文本和损坏内容输入至去噪扩散概率模型或者潜在扩散模型，通过 ControlNet 和 IPAdapter 将内容提示文本对应的文本提示特征向量确定为损坏内容在后向扩散过程的条件输入，从而生成与目标提示文本具有相似关系的第一处理内容 P2。例如计算机设备可以将第一图像内容 P0 中人物的嘴巴重绘为第一处理内容 P2 中微笑的嘴巴,通过对多个对象关键区域对应的目标提示文本(例如眼睛、鼻子、女孩等）对第一图像内容 P0 进行后向扩散，生成第一处理内容 P2。上述过程也可以称之为通过图像扩散模型对第一图像内容 P0 进行 InPainting(图像修复)。计算机设备可以将第一图像内容 P0 已替换为第一处理内容 P2 的图像，确定为问答保护图像 A2。

当目标内容保护处理方式为特征模糊重绘时，基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容的过程可以是：基于内容保护信息生成 K 个对象关键信息；K 为正整数，对象关键信息为用于识别得到内容保护信息的关键信息；对第一待处理区域中的图像内容与 K 个对象关键信息进行注意力处理，得到对象注意力结果向量，基于对象注意力结果向量在第一待处理区域中确定与 K 个对象关键信息相关联的 K 个对象关键区域，对 K 个对象关键区域对应的图像内容进行图像模糊重绘，得到 K 个已重绘图像内容，将包含 K 个已重绘图像内容的第一图像内容确定为第一处理内容。

具体的，计算机设备可以基于内容保护信息生成 K 个对象关键信息，对象关键信息可以是用于识别得到内容保护信息的关键信息。例如当内容保护信息为“人物/人脸”时，对象关键信息可以为“眼睛（虹膜）”、“鼻子”、“嘴唇（唇纹）”等。计算机设备可以对第一图像内容 P0 与 K 个对象关键信息进行注意力处理，得到对象注意力结果向量，基于对象注意力结果向量在第一待处理区域中确定与 K 个对象关键信息相关联的 K 个对象关键区域，例如可以得到每个对象关键信息在第一图像内容 P0 的图像区域，每个对象关键区域中的图像内容可以用于表示对象关键信息。

计算机设备可以对 K 个对象关键区域对应的图像内容进行图像模糊重绘，得到 K 个已重绘图像内容，将包含 K 个已重绘图像内容的第一图像内容 P0 确定为第一处理内容 P1。图像模糊重绘的方式可以包括高斯模糊、均值模糊、双边滤波等，本申请实施例在此不做限制。计算机设备可以将第一图像内容 P0 已替换为第一处理内容 P1 的图像，确定为问答保护图像 A1。

可以理解，计算机设备可以在图像中确定一个或多个第一待处理区域，每个第一待处理区域可以对应有不同的保护程度分值，即每个第一待处理区域的重绘方式可以互不相同。

步骤 S206，在问答注意力区域中确定第二待处理区域，将第二待处理区域中的图像内容确定为第二图像内容，对第二图像内容进行内容保护处理，得到第二处理内容；在图像中通过第一图像内容替换掉第一处理内容、且通过第二图像内容替换掉第二处理内容，得到包括第一处理内容和第二处理内容的问答保护图像。

具体的，由于问答注意力区域与内容保护区域之间存在重叠区域，问答注意力区域也保留有部分内容保护信息所指示的图像内容，为进一步提高问答注意力区域中内容保护信息所指示的图像内容的安全性，计算机设备可以也对问答注意力区域中的图像内容也进行内容保护处理。

计算机设备可以获取问答注意力区域对应的第一注意力分值，若第一注意力分值大于内容保护处理阈值（预先设置的数值，例如可以是 0.2），则计算机设备可以在问答注意力区域中确定第二待处理区域，若第一注意力分值小于或等于内容保护处理阈值，则计算机设备可以直接将问答注意力区域确定为第二待处理区域。计算机设备通过确定第二待处理区域，可以避开对图像中用于识别问题文本的关键图像内容进行内容保护处理，即可以在进一步提高内容保护信息的安全性的同时，依然可以保证图像问答的准确率。

计算机设备可以在问答注意力区域中确定第二待处理区域，其过程可以是：基于问题文本生成 T 个问题关键信息；T 为正整数，问题关键信息为用于识别得到问题文本的关键信息；对问答注意力区域中的图像内容与 T 个问题关键信息进行注意力处理，得到问题注意力结果向量，基于问题注意力结果向量在问答

注意力区域中确定与T个问题关键信息相关联的问题关键区域；将问答注意力区域中除问题关键区域之外的区域，确定为第二待处理区域。

具体的，计算机设备可以基于问题文本生成T个问题关键信息，问题关键信息可以是用于识别得到问题文本的关键信息，问题关键信息可以是问题文本中的词元（token）或词元的同义词，或者用于识别该词元的文本。问题关键信息可以表示图像中识别问题文本所对应的关键特征。例如，问题文本为“图像中的汽车为什么类型”，那么问题关键信息可以为 “汽车”、或者识别“汽车”的文本，可以包括“车轮”、“车门”、“车标”等。

计算机设备可以对问答注意力区域中的图像内容与T个问题关键信息进行注意力处理，得到问题注意力结果向量，基于问题注意力结果向量在问答注意力区域中确定与T个问题关键信息相关联的问题关键区域。每个问题关键区域中的图像内容可以用于表示问题文本。计算机设备可以将问答注意力区域中除问题关键区域之外的区域，确定为第二待处理区域。计算机设备可以获取问答注意力区域对应的问答注意力结果向量，基于问答注意力结果向量中与目标连通域相关联的第一注意力分值，生成问答注意力区域的相关程度分值，相关程度分值可以为0到1的分值，例如可以为0.8。计算机设备可以通过将相关程度分值设置为问答注意力区域的图像遮罩的灰度值来进行记录。

计算机设备可以基于相关程度分值确定问答注意力区域的保护程度分值，问答注意力区域的保护程度分值可以是数值1与问答注意力区域的相关程度分值之间的差值。例如，问答注意力区域的相关程度分值为0.8，则问答注意力区域的保护程度分值可以是0.2。各个保护程度分值所对应的内容保护处理方式可以参见上述步骤S205中的表1，在此不再赘述。

计算机设备可以基于问答注意力区域的保护程度分值确定问答注意力区域的目标内容保护处理方式，将第二待处理区域中的图像内容确定为第二图像内容，基于目标内容保护处理方式对第二图像内容进行内容保护处理，得到第二处理内容。目标内容保护处理方式的内容可以参见上述步骤S205的具体描述，在此不再赘述。其中，第二处理内容中被重绘的像素点可以包括与内容保护信息相关联的图像内容的像素点。

计算机设备可以在图像中通过第一图像内容替换掉第一处理内容、且通过第二图像内容替换掉第二处理内容，得到包括第一处理内容和第二处理内容的问答保护图像。

可以理解，由于问答注意力区域以及内容保护区域可以均为图像遮罩，因此第一待处理区域以及第二待处理区域也可以均为图像遮罩。图像遮罩可以用于隐藏或显示图像的特定部分。图像遮罩可以是一个与原图像大小相同的二值图像（即每个像素点只有0和1两种值），其中0代表透明（或隐藏），1代表不透明（或显示）。通过将图像遮罩与原图像结合，可以控制哪些部分的图像是可见的，哪些部分是被隐藏的。由于图像遮罩可以包含用于表征一个区域的图像顶点坐标，所以计算机设备可以通过图像遮罩中的图像顶点坐标来确定特定图像内容（如问答注意力区域中的图像内容或内容保护区域中的图像内容）的位置，基于内容保护区域对应的目标内容保护处理方式以及第二待处理区域的位置，对第二待处理区域中的图像内容进行局部重绘，以及基于问答注意力区域对应的目标内容保护处理方式以及问答注意力区域的位置，对问答注意力区域中的图像内容进行局部重绘，得到问答保护图像。问答保护图像可以用于生成问题文本对应的回答文本。其中，局部重绘可以是指通过图像遮罩，对第一待处理区域中的图像内容进行内容保护处理，将生成的新的图像内容直接覆盖或替换旧的图像内容，从而得到包含第一处理内容的问答保护图像。同样的，对第二待处理区域中的图像内容进行内容保护处理，将生成的新的图像内容直接覆盖或替换旧的图像内容，得到包含第一处理内容和第二图像内容的问答保护图像。局部重绘技术是一种图像处理方法，用于修改图像的特定区域而不影响其他部分。它包括选择需要修改的区域、生成新内容来替换旧内容、以及将新内容与原始图像无缝融合。这种方法常用于图像修复、对象移除和区域增强，通过复制粘贴、纹理合成或深度学习模型生成新内容，并使用混合算法确保自然过渡。

本申请实施例通过确定问答注意力区域和内容保护区域，可以区分模型处理图像问答所需的图像内容，进而基于问答注意力区域，在内容保护区域中确定第一待处理区域，对第一待处理区域中的第一图像内容进行内容保护处理，减少了模型在学习图像和回答问题文本的过程中，可能会导致与内容保护信息相关联的图像内容泄露的情况发生，同时，由于第一待处理区域与问答注意力区域互不重叠，所以模型依旧可以通过问答注意力区域识别到与问题文本相关联的图像内容，可见本申请可以在保证内容保护信息的安全性的同时，提高图像问答的准确率。通过内容保护区域确定保护程度分值，通过保护程度分值确定第一待处理区域的不同的内容保护处理方式，可以提高内容保护处理的多样性，降低内容保护信息泄露的风险。

另一方面，本申请实施例可以进一步提高问答注意力区域中图像内容的安全性，即也对问答注意力区域中的图像内容也进行内容保护处理。通过问答注意力区域确定相关程度分值，通过相关程度分值确定第

二待处理区域的不同的内容保护处理方式，可以对第一待处理区域中的图像内容采取不同的脱敏策略，遵循“最小必要信息暴露”原则，使得在保护内容保护信息的同时，尽可能不影响图像问答的准确率。

请参见图 6，图 6 是本申请实施例提供的一种数据处理的场景示意图三。如图 6 所示，对象（用户）可以在终端设备 100 中通过应用客户端上传图像以及图像对应的问题文本。其中，图像对应的问题文本可以是通过音频识别对象输入的语音得到的文本或者是对象在应用客户端直接输入的文本，本申请实施例在此不做限制。问题文本例如可以是“图像中的汽车的类型”。

终端设备 100 可以获取内容保护信息，内容保护信息可以是针对图像中的隐私的图像内容的描述信息，内容保护信息可以是一种类型描述或者直接指定的特定描述，可以是在应用客户端中的隐私内容类型里进行选择配置，也可以由对象输入，也可以是模型自动识别得到的。内容保护信息可以是文本数据、图像数据或者音频数据。内容保护信息的类型可以包括个人身份信息、联系信息、财务信息、车辆信息、医疗信息和地址位置信息等。内容保护信息例如可以是“人物/人像”。

可选的，当内容保护信息为图像数据时，可以是包含有指示隐私内容的图像框的图像，即用户可以在上传图像时，通过图像框框选需要进行内容保护处理的隐私内容，得到内容保护信息。

终端设备 100 中可以部署有端侧多模态模型。其中，端侧多模态模型可以是指用于在终端设备 100 上运行的，能够同时处理多种模态数据（如文本、图像、音频等）的人工智能模型，例如可以是 Phi-3-mini-128k-instruct 模型（一种开源轻量级模型），Chat-GLM（General Language Model）模型（一种自然语言生成模型）、Qwen-VLM（Vision Language Model）模型（一种大规模视觉语言模型）等可以在边缘设备上运行的多模态大语言模型。

终端设备 100 可以通过端侧多模态模型对图像与问题文本进行注意力处理，通过得到的问答注意力结果向量在图像中确定与问题文本相关联的问答注意力区域。同样的，终端设备 100 可以通过端侧多模态模型对图像与内容保护信息进行注意力处理，通过得到的内容保护结果向量在图像中确定与内容保护信息相关联的内容保护区域。问答注意力区域和内容保护区域均可以是一个通过多边形顶点描述的图像遮罩。

若问答注意力区域与内容保护区域之间存在重叠区域，则终端设备 100 可以将问答注意力区域与内容保护区域之间的重叠区域，确定为业务交叉区域，将内容保护区域中除业务交叉区域之外的区域，确定为第一待处理区域。终端设备 100 进一步保护问答注意力区域中与内容保护信息相关联的图像内容，终端设备 100 可以在问答注意力区域中确定为第二待处理区域。

终端设备 100 可以通过内容保护区域对应的内容保护结果向量确定保护程度分值，通过内容保护区域的保护程度分值确定针对第一待处理区域的内容保护处理的目标内容保护处理方式，终端设备 100 可以通过内容保护区域的目标内容保护处理方式对第一待处理区域进行内容保护处理，得到第一处理内容。

终端设备 100 可以通过问答注意力区域对应的问答注意力结果向量确定相关程度分值，通过问答注意力区域的相关程度分值确定针对第二待处理区域的内容保护处理的内容保护处理方式，终端设备 100 可以通过问答注意力区域的目标内容保护处理方式对第二待处理区域进行内容保护处理，得到第二处理内容。其中，内容保护处理方式可以包括图像隐身重绘、图像模糊重绘、同类替换重绘和特征模糊重绘等，本申请实施例在此不做限制。

终端设备 100 可以将第一图像内容已替换为第一处理内容、且第二图像内容已替换为第二处理内容的图像，确定为问答保护图像。将问答保护图像以及问题文本发送至业务服务器 200，业务服务器通过云端多模态模型对问答保护图像进行识别，生成问题文本对应的回答文本，将回答文本返回至终端设备 100。

本申请实施例通过确定问答注意力区域和内容保护区域，可以区分模型处理图像问答所需的图像内容，进而基于问答注意力区域，在内容保护区域中确定第一待处理区域，对第一待处理区域中的第一图像内容进行内容保护处理，减少了模型在学习图像和回答问题文本的过程中，可能会导致与内容保护信息相关联的图像内容泄露的情况发生，同时，由于第一待处理区域与问答注意力区域互不重叠，所以模型依旧可以通过问答注意力区域识别到与问题文本相关联的图像内容，可见本申请可以在保证图像问答的准确率同时，适当提高内容保护信息的安全性。通过内容保护区域确定保护程度分值，通过保护程度分值确定第一待处理区域的不同的内容保护处理方式，可以提高内容保护处理的多样性，降低内容保护信息泄露的风险。

另一方面，本申请实施例可以进一步提高问答注意力区域中图像内容的安全性，即也对问答注意力区域中的图像内容也进行内容保护处理。通过问答注意力区域确定相关程度分值，通过相关程度分值确定第二待处理区域的不同的内容保护处理方式，可以对第一待处理区域中的图像内容采取不用的脱敏策略，遵循“最小必要信息暴露”原则，使得在保护内容保护信息的同时，尽可能不影响图像问答的准确率。

请参见图 7，图 7 是本申请实施例提供的一种数据处理装置的结构示意图。如图 7 所示，该数据处理

装置1包括问答区域处理模块710、保护区域处理模块720和重绘区域处理模块730。

问答区域处理模块710，用于基于图像以及图像对应的问题文本，在图像中确定问答注意力区域；问答注意力区域中的图像内容为图像中与问题文本相关联的图像内容；

保护区域处理模块720，用于获取内容保护信息，基于图像以及内容保护信息，在图像中确定内容保护区域；内容保护区域中的图像内容为图像中与内容保护信息相关联的图像内容；

重绘区域处理模块730，用于若问答注意力区域与内容保护区域之间存在重叠区域，则基于问答注意力区域，在内容保护区域中确定第一待处理区域；第一待处理区域与问答注意力区域互不重叠；其中，第一待处理区域用于指示对图像中的第一图像内容进行内容保护处理，第一图像内容为图像中第一待处理区域中的图像内容。

在一种可能的实现方式中，重绘区域处理模块730用于基于问答注意力区域，在内容保护区域中确定第一待处理区域时，具体用于执行以下操作：

将问答注意力区域与内容保护区域之间的重叠区域，确定为业务交叉区域，将内容保护区域中除业务交叉区域之外的区域，确定为第一待处理区域。

在一种可能的实现方式中，重绘区域处理模块730还用于执行以下操作：

对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，在图像中通过第一处理内容替换掉第一图像内容，得到包括第一处理内容的问答保护图像；问答保护图像用于生成问题文本所对应的回答文本。

在一种可能的实现方式中，重绘区域处理模块730还用于执行以下操作：

对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容；

在问答注意力区域中确定第二待处理区域，将第二待处理区域中的图像内容确定为第二图像内容，对第二图像内容进行内容保护处理，得到第二处理内容；

在图像中通过第一图像内容替换掉第一处理内容、且通过第二图像内容替换掉第二处理内容，得到包括第一处理内容和第二处理内容的问答保护图像。

在一种可能的实现方式中，重绘区域处理模块730用于在所述问答注意力区域中确定第二待处理区域时，具体用于执行以下操作：

基于问题文本生成T个问题关键信息；T为正整数，问题关键信息为用于识别得到问题文本的关键信息；

对问答注意力区域中的图像内容与T个问题关键信息进行注意力处理，得到问题注意力结果向量，基于问题注意力结果向量在问答注意力区域中确定与T个问题关键信息相关联的问题关键区域；

将问答注意力区域中除问题关键区域之外的区域，确定为第二待处理区域。

在一种可能的实现方式中，问答区域处理模块710用于基于图像以及图像对应的问题文本，在图像中确定问答注意力区域时，具体用于执行以下操作：

对图像中的每个像素点进行连通域标记，得到M个连通域；M为正整数，连通域用于表示图像中像素点的像素值满足连通域标记条件且像素点的位置相邻的图像区域；

对M个连通域分别进行特征提取，得到图像特征向量，对图像对应的问题文本进行特征提取，得到问题文本特征向量；

对图像特征向量与问题文本特征向量进行注意力处理，得到问答注意力结果向量，基于问答注意力结果向量在图像中确定问答注意力区域。

在一种可能的实现方式中，问答区域处理模块710用于基于问答注意力结果向量在图像中确定问答注意力区域时，具体用于执行以下操作：

基于问答注意力结果向量生成M个连通域分别对应的注意力分值；

将具有最大注意力分值的连通域确定为目标连通域，基于目标连通域中像素点的位置信息生成边界轮廓点，将图像中边界轮廓点所包围的区域确定为问答注意力区域。

在一种可能的实现方式中，重绘区域处理模块730用于对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容时，具体用于执行以下操作：

获取内容保护区域相关联的保护程度分值；保护程度分值是基于内容保护信息对应的内容保护结果向量所生成的，内容保护结果向量是内容保护信息与图像进行注意力处理所得到的注意力向量；

基于保护程度分值确定目标重绘等级，在N个内容保护处理方式中确定目标重绘等级所指示的目标内容保护处理方式；N为正整数；

基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容。

在一种可能的实现方式中，目标内容保护处理方式为图像隐身重绘；重绘区域处理模块730用于基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容时，具体用于执行以下操作：

将图像中除第一待处理区域之外的图像内容，确定为重绘源内容；

获取重绘源内容中的背景图像，对背景图像进行特征提取，得到背景特征向量；

基于背景特征向量，对第一待处理区域中的图像内容进行重绘处理，得到第一处理内容；第一处理内容的内容特征向量与背景特征向量具有相似关系。

在一种可能的实现方式中，目标内容保护处理方式为图像模糊重绘；重绘区域处理模块730用于基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容时，具体用于执行以下操作：

获取模糊单元数值P，基于模糊单元数值P将第一待处理区域中的图像内容划分为P个待模糊图像单元；P为正整数；P个待模糊图像单元包括目标待模糊图像单元；

基于目标待模糊图像单元中每个像素点分别对应的颜色通道信息，生成目标待模糊图像单元的平均颜色通道信息，基于平均颜色通道信息对目标待模糊图像单元中的像素点进行模糊处理，得到已模糊图像单元；

当获取到P个待模糊图像单元分别对应的已模糊图像单元时，将P个已模糊图像单元确定为第一处理内容。

在一种可能的实现方式中，目标内容保护处理方式为同类替换重绘；重绘区域处理模块730用于基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容时，具体用于执行以下操作：

基于内容保护信息生成K个对象关键信息；K为正整数，对象关键信息为用于识别得到内容保护信息的关键信息；

对第一待处理区域中的图像内容与K个对象关键信息进行注意力处理，得到对象注意力结果向量，基于对象注意力结果向量在第一待处理区域中确定与K个对象关键信息相关联的对象关键区域；

获取与内容保护信息相关联的内容提示文本，在内容提示文本中确定与对象关键区域相关联的目标提示文本，对目标提示文本进行特征提取，得到文本提示特征向量；

对对象关键区域对应的图像内容分别进行损坏处理，得到损坏内容，将图像中除第一待处理区域之外的图像内容，确定为重绘源内容，对重绘源内容进行特征提取，得到重绘源特征向量；

基于文本提示特征向量和重绘源特征向量对损坏内容进行图像修复，得到第一处理内容。

在一种可能的实现方式中，目标内容保护处理方式为特征模糊重绘；重绘区域处理模块730用于基于目标内容保护处理方式，对第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容时，具体用于执行以下操作：

基于内容保护信息生成K个对象关键信息；K为正整数，对象关键信息为用于识别得到内容保护信息的关键信息；

对第一待处理区域中的图像内容与K个对象关键信息进行注意力处理，得到对象注意力结果向量，基于对象注意力结果向量在第一待处理区域中确定与K个对象关键信息相关联的对象关键区域，对对象关键区域对应的图像内容进行图像模糊重绘，得到第一处理内容。

本申请实施例通过确定问答注意力区域和内容保护区域，可以区分模型处理图像问答所需的图像内容，进而基于问答注意力区域，在内容保护区域中确定第一待处理区域，对第一待处理区域中的第一图像内容进行内容保护处理，减少了模型在学习图像和回答问题文本的过程中，可能会导致与内容保护信息相关联的图像内容泄露的情况发生，同时，由于第一待处理区域与问答注意力区域互不重叠，所以模型依旧可以通过问答注意力区域识别到与问题文本相关联的图像内容，可见本申请可以在保证图像问答的准确率同时，适当提高内容保护信息的安全性。通过内容保护区域确定保护程度分值，通过保护程度分值确定第一待处理区域的不同的内容保护处理方式，可以提高内容保护处理的多样性，降低内容保护信息泄露的风险。

另一方面，本申请实施例可以进一步提高问答注意力区域中图像内容的安全性，即也对问答注意力区域中的图像内容也进行内容保护处理。通过问答注意力区域确定相关程度分值，通过相关程度分值确定第二待处理区域的不同的内容保护处理方式，可以对第一待处理区域中的图像内容采取不用的脱敏策略，遵循“最小必要信息暴露”原则，使得在保护内容保护信息的同时，尽可能不影响图像问答的准确率。

本申请实施例中，术语“模块”或“单元”是指有预定功能的计算机程序或计算机程序的一部分，并与其他相关部分一起工作以实现预定目标，并且可以通过使用软件、硬件（如处理电路或存储器）或其组合来全部或部分实现。同样的，一个处理器（或多个处理器或存储器）可以用来实现一个或多个模块或单元。此外，每个模块或单元都可以是包含该模块或单元功能的整体模块或单元的一部分。

请参见图 8，图 8 是本申请实施例提供的一种计算机设备的结构示意图。如图 8 所示，该计算机设备 1000 可以包括：处理器 1001，网络接口 1004 和存储器 1005，此外，上述计算机设备 1000 还可以包括：用户接口 1003，和至少一个通信总线 1002。其中，通信总线 1002 用于实现这些组件之间的连接通信。其中，用户接口 1003 可以包括显示屏（Display）、键盘（Keyboard），可选用户接口 1003 还可以包括标准的有线接口、无线接口。网络接口 1004 可选的可以包括标准的有线接口、无线接口（如 WI-FI 接口）。存储器 1005 可以是高速 RAM 存储器，也可以是非不稳定的存储器（non-volatile memory），例如至少一个磁盘存储器。存储器 1005 可选的还可以是至少一个位于远离前述处理器 1001 的存储装置。如图 8 所示，作为一种计算机可读存储介质的存储器 1005 中可以包括操作系统、网络通信模块、用户接口模块以及设备控制应用程序。

在如图 8 所示的计算机设备 1000 中，网络接口 1004 可提供网络通讯网元；而用户接口 1003 主要用于为用户提供输入的接口；而处理器 1001 可以用于调用存储器 1005 中存储的设备控制应用程序，以实现：

基于图像以及图像对应的问题文本，在图像中确定问答注意力区域；问答注意力区域中的图像内容为图像中与问题文本相关联的图像内容；

获取内容保护信息，基于图像以及内容保护信息，在图像中确定内容保护区域；内容保护区域中的图像内容为图像中与内容保护信息相关联的图像内容；

若问答注意力区域与内容保护区域之间存在重叠区域，则基于问答注意力区域，在内容保护区域中确定第一待处理区域；第一待处理区域与问答注意力区域互不重叠；

其中，第一待处理区域用于指示对图像中的第一图像内容进行内容保护处理，第一图像内容为图像中第一待处理区域中的图像内容。

应当理解，本申请实施例中所描述的计算机设备 1000 可执行前文图 3 和图 4 任一个所对应实施例中对该数据处理方法的描述，在此不再赘述。另外，对采用相同方法的有益效果描述，也不再进行赘述。

此外，这里需要指出的是：本申请实施例还提供了一种计算机可读存储介质，且上述计算机可读存储介质中存储有计算机程序，当上述处理器执行上述计算机程序时，能够执行前文图 3 和图 4 任一个所对应实施例中对上述数据处理方法的描述，因此，这里将不再进行赘述。另外，对采用相同方法的有益效果描述，也不再进行赘述。对于本申请所涉及的计算机可读存储介质实施例中未披露的技术细节，请参照本申请方法实施例的描述。

上述计算机可读存储介质可以是前述任一实施例提供的数据处理装置或者上述计算机设备的内部存储单元，例如计算机设备的硬盘或内存。该计算机可读存储介质也可以是该计算机设备的外部存储设备，例如该计算机设备上配备的插接式硬盘，智能存储卡（smart media card，SMC），安全数字（secure digital，SD）卡，闪存卡（flash card）等。进一步地，该计算机可读存储介质还可以既包括该计算机设备的内部存储单元也包括外部存储设备。该计算机可读存储介质用于存储该计算机程序以及该计算机设备所需的其他程序和数据。该计算机可读存储介质还可以用于暂时地存储已经显示或者将要显示的数据。

此外，这里需要指出的是：本申请实施例还提供了一种计算机程序产品，该计算机程序产品包括计算机程序，该计算机程序存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机程序，处理器执行该计算机程序，使得该计算机设备执行前文图 3 和图 4 任一个所对应实施例提供的方法。

本申请实施例的说明书和权利要求书及附图中的术语“第一”、“第二”等是用于区别不同对象，而非用于描述特定顺序。此外，术语“包括”以及它们任何变形，意图在于覆盖不排他的包含。例如包含了一系列步骤或单元的过程、方法、装置、产品或设备没有限定于已列出的步骤或模块，而是可选地还包括没有列出的步骤或模块，或可选地还包括对于这些过程、方法、装置、产品或设备固有的其他步骤单元。

本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、计算机软件或者二者的结合来实现，为了清楚地说明硬件和软件的可互换性，在上述说明中已经按照网元一般性地描述了各示例的组成及步骤。这些网元究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的网元，但是这种实现不应认为超出本申请的范围。

本申请实施例提供的方法及相关装置是参照本申请实施例提供的方法流程图和/或结构示意图来描述的，具体可由计算机程序指令实现方法流程图和/或结构示意图的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。这些计算机程序指令可提供到通用计算机、专用计算机、嵌入式处理机或其他可编程设备的处理器以产生一个机器，使得通过计算机或其他可编程设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或结构示意图一个方框或多个方框中指定的功能的装置。这些计算机程序指令也可存储在能引导计算机或其他可编程设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和/或结构示意图一个方框或多个方框中指定的功能。这些计算机程序指令也可装载到计算机或其他可编程设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或结构示意一个方框或多个方框中指定的功能的步骤。

本申请实施例方法中的步骤可以根据实际需要进行顺序调整、合并和删减。

本申请实施例装置中的模块可以根据实际需要进行合并、划分和删减。

以上所揭露的仅为本申请较佳实施例而已，当然不能以此来限定本申请之权利范围，因此依本申请权利要求所作的等同变化，仍属本申请所涵盖的范围。

权利要求书

1、一种数据处理方法，其特征在于，包括：

基于图像以及所述图像对应的问题文本，在所述图像中确定问答注意力区域；所述问答注意力区域中的图像内容为所述图像中与所述问题文本相关联的图像内容；

获取内容保护信息，基于所述图像以及所述内容保护信息，在所述图像中确定内容保护区域；所述内容保护区域中的图像内容为所述图像中与所述内容保护信息相关联的图像内容；所述内容保护信息是预设的类型描述或者直接指定的特定描述；

若所述问答注意力区域与所述内容保护区域之间存在重叠区域，则基于所述问答注意力区域，在所述内容保护区域中确定第一待处理区域；所述第一待处理区域与所述问答注意力区域互不重叠；

其中，所述第一待处理区域用于指示对所述图像中的第一图像内容进行内容保护处理，所述第一图像内容为所述图像中所述第一待处理区域中的图像内容。

2、根据权利要求1所述的方法，其特征在于，所述基于所述问答注意力区域，在所述内容保护区域中确定第一待处理区域，包括：

将所述问答注意力区域与所述内容保护区域之间的重叠区域，确定为业务交叉区域，将所述内容保护区域中除所述业务交叉区域之外的区域，确定为第一待处理区域。

3、根据权利要求1所述的方法，其特征在于，还包括：

对所述第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，在所述图像中通过所述第一处理内容替换掉所述第一图像内容，得到包括所述第一处理内容的问答保护图像；所述问答保护图像用于生成所述问题文本所对应的回答文本。

4、根据权利要求1所述的方法，其特征在于，还包括：

对所述第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容；

在所述问答注意力区域中确定第二待处理区域，将所述第二待处理区域中的图像内容确定为第二图像内容，对所述第二图像内容进行内容保护处理，得到第二处理内容；

在所述图像中通过所述第一图像内容替换掉所述第一处理内容、且通过所述第二图像内容替换掉所述第二处理内容，得到包括所述第一处理内容和所述第二处理内容的问答保护图像。

5、根据权利要求4所述的方法，其特征在于，所述在所述问答注意力区域中确定第二待处理区域，包括：

基于所述问题文本生成T个问题关键信息；T为正整数，所述问题关键信息为用于识别得到所述问题文本的关键信息；

对所述问答注意力区域中的图像内容与所述T个问题关键信息进行注意力处理，得到问题注意力结果向量，基于所述问题注意力结果向量在所述问答注意力区域中确定与所述T个问题关键信息相关联的问题关键区域；

将所述问答注意力区域中除所述问题关键区域之外的区域，确定为第二待处理区域。

6、根据权利要求1所述的方法，其特征在于，所述基于图像以及所述图像对应的问题文本，在所述图像中确定问答注意力区域，包括：

对图像中的每个像素点进行连通域标记，得到M个连通域；M为正整数，所述连通域用于表示所述图像中像素点的像素值满足连通域标记条件且像素点的位置相邻的图像区域；

对所述M个连通域分别进行特征提取，得到图像特征向量，对所述图像对应的问题文本进行特征提取，得到问题文本特征向量；

对所述图像特征向量与所述问题文本特征向量进行注意力处理，得到问答注意力结果向量，基于所述问答注意力结果向量在所述图像中确定问答注意力区域。

7、根据权利要求6所述的方法，其特征在于，所述基于所述问答注意力结果向量在所述图像中确定问答注意力区域，包括：

基于所述问答注意力结果向量生成所述M个连通域分别对应的注意力分值；

将具有最大注意力分值的连通域确定为目标连通域，基于所述目标连通域中像素点的位置信息生成边界轮廓点，将所述图像中所述边界轮廓点所包围的区域确定为问答注意力区域。

8、根据权利要求3所述的方法，其特征在于，所述对所述第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，包括：

获取所述内容保护区域相关联的保护程度分值；所述保护程度分值是基于所述内容保护信息对应的内

容保护结果向量所生成的，所述内容保护结果向量是所述内容保护信息与所述图像进行注意力处理所得到的注意力向量；

基于所述保护程度分值确定目标重绘等级，在N个内容保护处理方式中确定所述目标重绘等级所指示的目标内容保护处理方式；N为正整数；

基于所述目标内容保护处理方式，对所述第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容。

9、根据权利要求8所述的方法，其特征在于，所述目标内容保护处理方式为图像隐身重绘；所述基于所述目标内容保护处理方式，对所述第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，包括：

将所述图像中除所述第一待处理区域之外的图像内容，确定为重绘源内容；

获取所述重绘源内容中的背景图像，对所述背景图像进行特征提取，得到背景特征向量；

基于所述背景特征向量，对所述第一待处理区域中的图像内容进行重绘处理，得到第一处理内容；所述第一处理内容的内容特征向量与所述背景特征向量具有相似关系。

10、根据权利要求8所述的方法，其特征在于，所述目标内容保护处理方式为图像模糊重绘；所述基于所述目标内容保护处理方式，对所述第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，包括：

获取模糊单元数值P，基于所述模糊单元数值P将所述第一待处理区域中的图像内容划分为P个待模糊图像单元；P为正整数；所述P个待模糊图像单元包括目标待模糊图像单元；

基于所述目标待模糊图像单元中每个像素点分别对应的颜色通道信息，生成所述目标待模糊图像单元的平均颜色通道信息，基于所述平均颜色通道信息对所述目标待模糊图像单元中的像素点进行模糊处理，得到已模糊图像单元；

当获取到所述P个待模糊图像单元分别对应的已模糊图像单元时，将P个已模糊图像单元确定为第一处理内容。

11、根据权利要求8所述的方法，其特征在于，所述目标内容保护处理方式为同类替换重绘；所述基于所述目标内容保护处理方式，对所述第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，包括：

基于所述内容保护信息生成K个对象关键信息；K为正整数，所述对象关键信息为用于识别得到所述内容保护信息的关键信息；

对所述第一待处理区域中的图像内容与所述K个对象关键信息进行注意力处理，得到对象注意力结果向量，基于所述对象注意力结果向量在所述第一待处理区域中确定与所述K个对象关键信息相关联的对象关键区域；

获取与所述内容保护信息相关联的内容提示文本，在所述内容提示文本中确定与所述对象关键区域相关联的目标提示文本，对所述目标提示文本进行特征提取，得到文本提示特征向量；

对所述对象关键区域对应的图像内容分别进行损坏处理，得到损坏内容，将所述图像中除所述第一待处理区域之外的图像内容，确定为重绘源内容，对所述重绘源内容进行特征提取，得到重绘源特征向量；

基于所述文本提示特征向量和所述重绘源特征向量对所述损坏内容进行图像修复，得到第一处理内容。

12、根据权利要求8所述的方法，其特征在于，所述目标内容保护处理方式为特征模糊重绘；所述基于所述目标内容保护处理方式，对所述第一待处理区域中的图像内容进行内容保护处理，得到第一处理内容，包括：

基于所述内容保护信息生成K个对象关键信息；K为正整数，所述对象关键信息为用于识别得到所述内容保护信息的关键信息；

对所述第一待处理区域中的图像内容与所述K个对象关键信息进行注意力处理，得到对象注意力结果向量，基于所述对象注意力结果向量在所述第一待处理区域中确定与所述K个对象关键信息相关联的对象关键区域，对所述对象关键区域对应的图像内容进行图像模糊重绘，得到第一处理内容。

13、一种数据处理装置，其特征在于，包括：

问答区域处理模块，用于基于图像以及所述图像对应的问题文本，在所述图像中确定问答注意力区域；所述问答注意力区域中的图像内容为所述图像中与所述问题文本相关联的图像内容；

保护区域处理模块，用于获取内容保护信息，基于所述图像以及所述内容保护信息，在所述图像中确定内容保护区域；所述内容保护区域中的图像内容为所述图像中与所述内容保护信息相关联的图像内容；

所述内容保护信息是预设的类型描述或者直接指定的特定描述；

重绘区域处理模块，用于若所述问答注意力区域与所述内容保护区域之间存在重叠区域，则基于所述问答注意力区域，在所述内容保护区域中确定第一待处理区域；所述第一待处理区域与所述问答注意力区域互不重叠；其中，所述第一待处理区域用于指示对所述图像中的第一图像内容进行内容保护处理，所述第一图像内容为所述图像中所述第一待处理区域中的图像内容。

14、一种计算机设备，其特征在于，包括：处理器、存储器以及网络接口；

所述处理器与所述存储器、所述网络接口相连，其中，所述网络接口用于提供数据通信功能，所述存储器用于存储计算机程序，所述处理器用于调用所述计算机程序，以使得所述计算机设备执行权利要求 1-12 任一项所述的方法。

15、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质中存储有计算机程序，所述计算机程序适于由处理器加载并执行，以使得具有所述处理器的计算机设备执行权利要求 1-12 任一项所述的方法。

16、一种计算机程序产品，其特征在于，所述计算机程序产品包括计算机程序，所述计算机程序存储在计算机可读存储介质中，且适于由处理器读取并执行，以使得具有所述处理器的计算机设备执行权利要求1-12任一项所述的方法。