

背景图像生成方法、电子设备及计算机存储介质

本申请要求申请日为 2023 年 10 月 18 日、申请号为“202311355147.8”、专利名称为“背景图像生成方法、电子设备及计算机存储介质”的中国发明专利申请的优先权，其全部内容在此引入作为参考。

技术领域

本申请实施例涉及计算机技术领域，尤其涉及一种背景图像生成方法、电子设备及计算机存储介质。

背景技术

随着计算机技术的发展，在人们的日常生活和工作中经常有需要电子扫描文件的情况，例如，书籍电子化、发票报销、工作文档扫描打印、证件材料扫描件等等应用场景。因专业的扫描仪价格昂贵，扫描软件应运而生。通过扫描软件，用户可以随时随地实现方便、低成本的文件扫描。伴随着扫描软件的广泛使用，扫描场景也越来越多样化，这就要求扫描软件能够不断地迭代更新，以适应新场景需求。但扫描软件的迭代更新，需要以大量的各种扫描场景下的训练样本为前提。为此，在现有的一种方式中，利用专业的渲染软件（如 unity3D 等）搭建虚拟的光照环境，其可以包含不同的光源、不同的遮挡物、不同的纹理背景等，通过变换光源的位置、强弱、色温等再搭配不同的遮挡物以及纸面纹理来模拟用户扫描文件的场景，以使生成的扫描图像的背景更为逼真，进而使得扫描图像较为逼真。但这种方式中，一是这类专业渲染软件使用门槛较高，需要长时间的调试与学习才能达到一定的模拟效果，时间成本也较大，造成整体实现成本较高；二是这类专业渲染软件模拟生成的图像背景在数据分布上和真实的样本分布存在比较大的偏差，使得生成的扫描图像的背景效果较差，进而导致虽然实现软件扫描的机器学习模型在模拟的训练样本上效果很好，但在真实数据上却效果很差。

发明内容

有鉴于此，本申请实施例提供一种背景图像生成方案，以至少部分解决上述问题。根据本申请实施例的第一方面，提供了一种背景图像生成方法，包括：获取基于物理文件生成的真实图像；对所述真实图像进行多种不同类型的目标内容检测，获得对应的多个目标内容图像区域及对应的多个目标内容掩膜；对所述多个目标内容掩膜进行融合，获得融合掩膜；从所述真实图像中，擦除所述融合掩膜对应的图像区域，并根据擦除结果获得所述真实图像对应的背景图像。

根据本申请实施例的第二方面，提供了一种电子设备，包括：处理器、存储器、通信接口和通信总线，所述处理器、所述存储器和所述通信接口通过所述通信总线完成相互间的通信；所述存储器用于存放至少一可执行指令，所述可执行指令使所述处理器执

行如第一方面所述方法对应的操作。

根据本申请实施例的第三方面，提供了一种计算机存储介质，其上存储有计算机程序，该程序被处理器执行时实现如第一方面所述的方法。

根据本申请实施例的再一方面，提供了一种计算机程序产品，包括计算机指令，该计算机指令指示计算设备执行如第一方面所述的方法对应的操作。

根据本申请实施例提供的方案，可以基于真实的包含文件信息的图像，对其进行目标内容检测和掩膜处理，并基于多个目标内容掩膜生成融合掩膜，基于融合掩膜进行图像区域擦除，从而获得真实图像对应的背景图像。由此，一方面，无需图像处理人员进行复杂的参数学习和调试，基于真实图像即可生成相应的背景图像，降低了背景图像生成成本；另一方面，背景图像基于真实图像生成，因而在数据分布上和真实的样本分布一致，不仅能够使得生成的扫描图像的背景效果较好；并且，后续，基于背景图像来合成扫描图像以作为训练样本时，也能够使得实现软件扫描的机器学习模型的训练效果更好。

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实施例中记载的一些实施例，对于本领域普通技术人员来讲，还可以根据这些附图获得其他的附图。

- 图 1 为适用本申请实施例方案的示例性系统的示意图。
- 图 2 为根据本申请实施例的一种背景图像生成方法的步骤流程图。
- 图 3 为本申请的步骤 S104 的一个可选的子步骤流程图。
- 图 4 为本申请的步骤 S106 的一个可选的子步骤流程图。
- 图 5 为本申请的步骤 S108 的一个可选的子步骤流程图。
- 图 6 为本申请实施例中的一种场景示例的示意图。
- 图 7A 为本申请实施例中的合成扫描文件样本图像的一种场景示例的示意图。
- 图 7B 为本申请实施例中的合成扫描文件样本图像的另一种场景示例的示意图。
- 图 8 为根据本申请实施例的一种电子设备的结构示意图。

具体实施方式

为了使本领域的人员更好地理解本申请实施例中的技术方案，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，所描述的实施例仅是本申请实施例一部分实施例，而不是全部的实施例。基于本申请实施例中的实施例，本领域普通技术人员所获得的所有其他实施例，都应当属于本申请实施例保护的

范围。

下面结合本申请实施例附图进一步说明本申请实施例具体实现。

图 1 示出了一种适用本申请实施例方案的示例性系统。如图 1 所示，该系统 100 可

以包括云服务器端 102、通信网络 104 和/或一个或多个用户设备 106，图 1 中示例为多个用户设备。需要说明的是，本申请实施例的方案可以由云服务器端 102 和用户设备 106 协作完成，也可以由用户设备 106 独立完成。

云服务器端 102 可以是用于存储信息、数据、程序和/或任何其他合适类型的内容的任何适当的设备，包括但不限于分布式存储系统设备、服务器集群、计算云服务器端集群等。在一些实施例中，云服务器端 102 可以执行任何适当的功能。例如，当云服务器端 102 和用户设备 106 协作完成本申请实施例的方案时，在一些实施例中，云服务器端 102 可以接收用户设备 106 发送来的基于物理文件生成的真实图像，生成真实图像对应的背景图像。作为可选的示例，在一些实施例中，云服务器端 102 可以先对真实图像进行多种不同类型的目标内容检测，获得对应的多个目标内容图像区域及对应的多个目标内容掩膜；进而，可对多个目标内容掩膜进行融合，获得融合掩膜；再从真实图像中，擦除融合掩膜对应的图像区域，并根据擦除结果获得真实图像对应的背景图像。作为另一示例，在一些可选的实施例中，云服务器端 102 可以将生成的背景图像发送到用户设备 106。在一些可选的实施例中，云服务器端 102 可以基于背景图像合成扫描文件样本图像（例如基于背景图像和预设的文件内容图像，合成扫描文件样本图像；或者，基于背景图像和预设的文件内容文本，合成扫描文件样本图像），并基于扫描文件样本图像对用于生成扫描文件的机器学习模型进行训练。

在一些实施例中，通信网络 104 可以是一个或多个有线和/或无线网络的任何适当的组合。例如，通信网络 104 能够包括以下各项中的任何一种或多种：互联网、内联网、广域网(Wide Area Network, WAN)、局域网(Local Area Network, LAN)、无线网络、数字订户线路(Digital Subscriber Line, DSL)网络、帧中继网络、异步转移模式(Asynchronous Transfer Mode, ATM)网络、虚拟专用网(Virtual Private Network, VPN)和/或任何其它合适的通信网络。用户设备 106 能够通过一个或多个通信链路(例如，通信链路 112)连接到通信网络 104，该通信网络 104 能够经由一个或多个通信链路(例如，通信链路 114)被链接到云服务器端 102。通信链路可以是适合于在用户设备 106 和云服务器端 102 之间传送数据的任何通信链路，诸如网络链路、拨号链路、无线链路、硬连线链路、任何其它合适的通信链路或此类链路的任何合适的组合。

用户设备 106 可以包括适合于呈现图像、与用户交互等的任何一个或多个用户设备。当由用户设备 106 独立完成本申请实施例的方案时，用户设备 106 可以用于基于物理文件生成的真实图像生成真实图像对应的背景图像。作为可选的示例，在一些实施例中，用户设备 106 可以先对真实图像进行多种不同类型的目标内容检测，获得对应的多个目标内容图像区域及对应的多个目标内容掩膜；进而，可对多个目标内容掩膜进行融合，获得融合掩膜；再从真实图像中，擦除融合掩膜对应的图像区域，并根据擦除结果获得真实图像对应的背景图像。在一些可选的实施例中，用户设备 106 可以基于背景图像合成扫描文件样本图像（例如基于背景图像和预设的文件内容图像，合成扫描文件样本图像；或者，基于背景图像和预设的文件内容文本，合成扫描文件样本图像）。进而，

用户设备 106 可以将扫描文件样本图像发送至云服务器端 102，以对云服务器端 102 中的用于生成扫描文件的机器学习模型进行训练。在一些实施例中，用户设备 106 可以包括任何合适类型的设备。例如，在一些实施例中，用户设备 106 可以包括移动设备、平板电脑、膝上型计算机、台式计算机和/或任何其他合适类型的用户设备。

- 5 基于上述系统，本申请实施例提供了一种背景图像生成方案，以下通过多个实施例进行说明。

图 2 为根据本申请实施例的一种背景图像生成方法的步骤流程图。根据本申请中的第一方面，提供了一种背景图像生成方法，参照图 2 所示，该方法包括步骤 S102、S104、S106 和 S108，具体地：

- 10 S102：获取基于物理文件生成的真实图像。

本申请中，物理文件可以是指用于扫描的实物文件。物理文件可以为任意类型，例如，物理文件包括但不限于可以是书籍、发票、合同、营业执照、工作文档、证件材料、身份证、驾驶证等。

- 15 本申请中的真实图像，具有真实样本的特征，其数据分布上与真实的样本一致，这使得基于真实图像生成背景图像，得到的背景图像在数据分布上和真实的样本分布一致。

- 20 可选地，本申请中步骤 S102 中的基于物理文件生成的真实图像，可以是预先通过摄像头拍摄物理文件得到的图像。例如，以物理文件为发票进行举例，则基于发票生成的真实图像可以以如下方式得到：用户将需要扫描的发票置于支撑物（例如桌面等）上或者持于手中，然后使用移动终端上的摄像头从朝向发票的方向拍摄发票，从而得到相应的真实图像。应理解，这只是一个示例，并不作为对本申请中的任何限制。

应理解，步骤 S102 中获取基于物理文件生成的真实图像，可以是在有需要的时候进行拍摄得到真实图像然后直接获取，也可以是事先进行拍摄得到真实图像，然后将真实图像存储到一个预定的存储空间（包括但不限于磁盘、硬盘、存储器等存储介质，也可以是数据库等）中，在需要时直接从存储空间中获取，对此本申请中不进行任何限制。

- 25 可选地，真实图像也可以包括模型处理后的存在瑕疵的图像（例如，这些图像可以是来自于用户（可以包括线上用户和/或离线用户）使用扫描软件的历史过程中，被扫描软件的用于实现软件扫描的机器学习模型处理后存在瑕疵的图像不断积累得来），可以由人工从模型处理后的存在瑕疵的图像挑选出需要的真实图像后存储到存储空间中，或者，也可以通过算法按照规则，从模型处理后的存在瑕疵的图像中捞取出需要的真实图像后存储到存储空间中。例如，瑕疵可以是图像中有包括但不限于阴影、折痕、透字、手指、夹子等目标，这些目标可作为真实图像的背景的一部分，属于真实样本的特征，其数据分布上与真实的样本一致。通过这样的真实图像来生成背景图像，并在后续使用背景图像合成扫描文件样本图像，以对用于生成扫描文件的机器学习模型进行训练，可以扩展机器学习模型的能力边界。

- 30 S104：对真实图像进行多种不同类型的目标内容检测，获得对应的多个目标内容图像区域及对应的多个目标内容掩膜。

步骤 S102 中获取的基于物理文件生成的真实图像中,包括有多种不同类型的目标内容,这些目标内容通常为有实际意义的信息内容,因此,本申请步骤 S104 中可以通过对真实图像进行多种不同类型的目标内容检测,获得对应的多个目标内容图像区域及对应的多个目标内容掩膜(mask),以便于后续步骤利用多个目标内容掩膜进行数据处理,以得到更能满足需求的背景图像。

本申请中,目标内容可以视相应的真实图像确定。在一些可选的实施例,本申请中的目标内容包括以下类型至少之一:文字类型内容、图表类型内容、色块图案类型内容、印章类型内容。可选地,参照图 3 所示的流程图,步骤 S104 包括子步骤 S1041、S1042 和 S1043,具体地:

S1041: 针对真实图像进行以下目标内容检测至少之二: 针对文字的检测、针对图表的检测、针对色块图案的检测、针对印章的检测。

本申请中,可以采用任意可行的目标检测算法实现针对文字、图表、色块图案和印章的检测。

本申请中通过对真实图像进行文字、图表、色块图案、印章中至少二种不同类型的目标内容检测,使得本申请中的背景图像生成方法,能便于更好适应真实生活中进行文件扫描的需求。

此外,本申请中不对文字、图表、色块图案、印章进行具体限定。例如可参照如下说明进行示意性理解,文字可以包括但不限于汉字、阿拉伯数字、英文文字等;图表可以包括但不限于柱状图、饼状图、折线图、表格等;色块图案可以包括任意的规则色块图案或者不规则色块图案;印章可以可以包括但不限于圆形印章、矩形印章等规则形状印章,也可以是别的形状的印章。

还需要说明的是,虽然本申请中是对真实图像进行文字、图表、色块图案、印章中至少二种不同类型的目标内容检测,但这并不意味着一定能从真实图像中检测出多种不同类型的目标内容,这取决于目标图像的实际情况,本申请中对此不进行任何限制。例如,如果真实图像中只有文字类型内容,而无图表、色块图案、印章等类型内容,则虽然会进行多种不同类型的目标检测,但得到的检测结果中可能只有文字类型内容对应的结果。当然,这只是一个示例,并不作为对本申请中的任何限制。

此外,还需要说明的是,本申请中,“多种”、“多个”等与“多”有关的数量意指两个及两个以上。

S1042: 根据检测结果,获得检测出的多种类型的目标内容对应的多个图像区域。

具体地,可以根据子步骤 S1041 中得到的目标内容检测的检测结果,从真实图像中确定检测出的多种类型的目标内容对应的多个图像区域,这些图像区域包括了真实图像中至少部分除背景以外的内容。

由于子步骤 S1041 中是对真实图像进行文字、图表、色块图案、印章中至少二种不同类型的目标内容检测,因此在检测出相应的内容后,可以对其进行合适的处理以得到相应的图像区域。例如,在一些可选的实施例,子步骤 S1042,包括以下子步骤 S1042A、

S1042B、S1042C 和 S1042D 中的至少之一，具体地：

S1042A：若根据检测结果，确定检测出文字类型内容，则对文字类型内容进行文字识别，以根据文字识别结果，确定各个文字分别对应的图像区域。

5 可选地，可以采用任意合适的文字识别算法来对文字类型内容进行识别，以根据文字识别结果，确定各个文字分别对应的图像区域。例如，文字识别算法包括但不限于 OCR（Optical Character Recognition，光学字符识别）算法等。

S1042B：若根据检测结果，确定检测出图表类型内容，则对图表类型内容进行图像分割，以获得图表类型内容对应的图像区域。

10 可选地，可以采用任意合适的图像分割算法来对图表类型内容进行图像分割，以获得图表类型内容对应的图像区域。例如，图像分割算法包括但不限于语义分割算法、实例分割算法等等。

S1042C：若根据检测结果，确定检测出色块图案类型内容，则对色块图案类型内容进行图像分割，以获得色块图案内容对应的图像区域。

15 可选地，可以采用任意合适的图像分割算法来对色块图案类型内容进行图像分割，以获得色块图案类型内容对应的图像区域。例如，图像分割算法包括但不限于语义分割算法、实例分割算法等等。

S1042D：若根据检测结果，确定检测出印章类型内容，则对印章类型内容进行文字识别及边框识别，以根据识别结果，确定印章对应的图像区域。

20 可选地，可以采用任意合适的文字识别算法和任意合适的边框识别算法来对印章类型内容进行识别，以获得印章类型内容对应的图像区域。例如，其中的文字识别算法包括但不限于 OCR（Optical Character Recognition，光学字符识别）算法等。

基于此，本申请中通过子步骤 S1042A~S1042D 的可选实施方式，对于根据检测结果确定出的不同的目标内容，采用不同的方式来适应性地确定目标内容对应的图像区域，可以准确有效地得到图像区域的确定结果。

25 S1043：对多个图像区域进行掩膜，获得对应的多个目标内容掩膜。

在子步骤 S1042 中得到了多个目标区域之后，可以对多个图像区域进行掩膜处理，以获得对应的多个目标内容掩膜（mask）。其中，掩膜的具体获得方式可以由本领域技术人员采用任意适当方式实现，包括但不限于通过用于对图像中的目标对象进行掩模的机器学习模型来获得的方式。此外，可选地，该机器学习模型在输出图像区域对应的掩膜时，还可同时输出各个掩膜对应的置信度。

基于此，本申请中通过上述子步骤 S1041~S1043 的可选实施方式，可以准确有效地实现对真实图像进行多种不同类型的目标内容检测，以获得对应的多个目标内容图像区域及对应的多个目标内容掩膜，从而便于后续利用其进行数据处理以得到更能满足需求的背景图像。

35 S106：对多个目标内容掩膜进行融合，获得融合掩膜。

在得到多个目标内容掩膜后，将多个目标内容掩膜进行融合（如通过合并的方式进

行融合等)，得到更能满足数据处理需求的融合掩膜，通过融合掩膜，一方面可以更好地提高后续步骤生成背景图像的效果，另一方面也便于降低后续生成背景图像时的计算量。

- 本申请中不限定步骤 S106 的具体实现，能够满足需要即可。在一些可选的实施例中，
- 5 参照图 4 所示的流程图，步骤 S106 包括子步骤 S1061 和 S1062，具体地：

S1061：根据多个目标内容掩膜中各目标内容掩膜的置信度，对多个目标内容掩膜进行过滤。

- 例如，如前所述，机器学习模型在输出图像区域对应的掩膜时，还可没输出掩膜对应的置信度，则本步骤中，可以根据多个目标内容掩膜中各目标内容掩膜的置信度，将
- 10 置信度较低的目标内容掩膜过滤掉，以保证目标内容掩膜的准确性。

- 本申请中不限定子步骤 S1061 的具体实现。在一些可选的实施例中，可以通过设置置信度阈值的方式进行过滤，例如，步骤 S1061 具体可以通过判断多个目标内容掩膜中的各个目标内容掩膜的置信度是否大于置信度阈值，若是，则保留，若否，则不保留，从而实现对多个目标内容掩膜进行过滤的目的。置信度阈值可以按照需要设置，例如可
- 15 以设置为 70%、80%等等。

或者，子步骤 S1061 也可以按照需要设置其他的条件进行过滤，本申请在此不进行具体限制。

S1062：对过滤后的目标内容掩膜进行融合，获得融合掩膜。

示例性地，可以采用诸如掩膜合并等方式对目标内容掩膜进行融合。

- 基于此，本申请中通过上述子步骤 S1061~S1062 的可选实施方式，可以有效地对多个目标内容掩膜进行融合获得融合掩膜，一方面，将多个目标内容掩膜融合为一个整体，可以更便于后续处理，如擦除处理，提高处理速度；另一方面，也可以更好地避免后续擦除融合掩膜对应的图像区域时将真实图像中的过多的背景区域擦除，从而更好地提高后续步骤生成背景图像的效果以及降低后续生成背景图像时的计算量。
- 20

- S108：从真实图像中，擦除融合掩膜对应的图像区域，并根据擦除结果获得真实图像对应的背景图像。
- 25

在步骤 S106 中得到融合掩膜后，可从真实图像中擦除融合掩膜对应的图像区域，使得擦除后的真实图像只剩背景，从而可以根据擦除结果获得真实图像对应的背景图像。

- 基于此，本申请中通过上述步骤 S102~S108 的背景图像生成方法，可以基于真实的包含文件信息的图像，对其进行目标内容检测和掩膜处理，并基于多个目标内容掩膜生成融合掩膜，基于融合掩膜进行图像区域擦除，从而获得真实图像对应的背景图像。由此，一方面，无需图像处理人员进行复杂的参数学习和调试，基于真实图像即可生成相应的背景图像，降低了背景图像生成成本；另一方面，背景图像基于真实图像生成，因而在数据分布上和真实的样本分布一致，不仅能够使得生成的扫描图像的背景效果较好；
- 30
- 并且，后续，基于背景图像来合成扫描图像以作为训练样本时，也能够使得实现软件扫描的机器学习模型的训练效果更好。
- 35

本申请中，不限定步骤 S108 的具体实现，只要能够实现相应的需求即可。例如，在实现步骤 S108 及其各可选实施方式时，可以从真实图像中将融合掩膜对应的图像区域的内容进行擦除得到擦除结果，并用临近区域内容对擦除区域进行填充，以根据擦除结果获得真实图像对应的背景图像。例如，擦除可以采用包括但不限于传统的图像填充技术、

5 基于人工智能（AI）的 image_inpainting 技术来实现，也可以采用比的方式实现，本申请对此不作任何限定。

在一些可选的实施例中，在步骤 S106 中获得融合掩膜之后，该背景图像生成方法还包括：根据形态学处理参数，对融合掩膜进行形态学处理；步骤 S108 中的“从真实图像中，擦除融合掩膜对应的图像区域”包括：从真实图像中，擦除进行了形态学处理的融合掩膜对应的图像区域。

10

可选地，通过对融合掩膜进行形态学处理，并从真实图像中擦除进行了形态学处理的融合掩膜，根据这样的擦除结果获得真实图像对应的背景图像的效果更好。

可选地，对融合掩膜进行形态学处理，可以包括腐蚀处理和膨胀处理中的至少一者。可以按照不同的真实图像的需要，采用不同的形态学参数，进行不同类型的形态学处理，

15 本申请中对此不进行特别限定。

在一些可选的实施例中，参照图 5 所示的流程图，步骤 S108 包括子步骤 S1081 和 S1082，具体地：

S1081：从真实图像中，擦除融合掩膜对应的图像区域，获得擦除后图像；

S1082：计算擦除后图像的信息熵，若信息熵满足预设信息熵阈值，则将擦除后图像

20 作为真实图像对应的背景图像。

具体地，子步骤 S1082 中，信息熵可以用来表示图像中内容的丰富程度，信息熵的值越高则表示信息量越高，内容越丰富，反之则内容越少。由此，可以通过计算得到擦除后图像的信息熵来判断真实图像中的融合掩膜对应的图像区域是否擦除干净。

本申请中可以通过预设信息熵阈值来进行判断擦除后图像的信息熵是否满足要求，

25 若擦除后图像的信息熵满足预设信息熵阈值，则可以认为真实图像中的融合掩膜对应的图像区域已擦除干净，可以作为真实图像对应的背景图像使用。

可选地，如果擦除后图像的信息熵低于预设信息熵阈值，则擦除后图像的信息熵满足预设信息熵阈值，反之，若擦除后图像的信息熵高于预设信息熵阈值，则擦除后图像的信息熵不满足预设信息熵阈值。在擦除后图像的信息熵满足预设信息熵阈值，则可以

30 将擦除后图像作为真实图像对应的背景图像。

基于此，本申请中通过子步骤 S1081~S1082 的可选实施方式，可以保证真实图像中的融合掩膜对应的图像区域擦除干净，将擦除干净得到的擦除后图像作为真实图像的背景图像使用，能够使得生成的扫描图像的背景效果更好，并且后续基于背景图像来合成扫描图像以作为训练样本时，也能够使得实现软件扫描的机器学习模型的训练效果更好。

35 本申请中，预设信息熵阈值可以按照实际需要设定。在一些可选的实施例中，预设信息熵阈值通过以下方式确定：确定与当前处理的真实图像相似的预设数量的历史图像，

根据历史图像对应的背景图像的信息熵中的最小值，确定预设信息熵阈值。通过这样的方式，可以有更合理地确定预设信息熵阈值，使得确定擦除后图像的信息熵是否满足预设信息熵阈值的结果更合理。

具体地，历史图像也即历史上的真实图像，与当前正在处理的真实图像只是时间上的区别。可选地，本申请中所说的历史图像与当前处理的真实图像相似，可以是指历史图像的类型和内容布局中的至少之一近似。例如，历史图像的类型是基于发票生成的真实图像，当前处理的真实图像也是基于发票生成的真实图像，则认为两者相似；又例如，历史图像的类型是基于营业执照生成的真实图像，当前处理的真实图像也是基于营业执照生成的真实图像，则认为两者相似；再例如，历史图像是基于一本书籍的具有文字类型内容和图表类型内容的某一页生成的真实图像，当前处理的真实图像是基于合同的具有文字类型内容和图表类型内容的某一页生成的真实图像，则也认为两者相似。应理解，这几个例子仅仅是便于理解的示例，并不作为对本申请的任何限制。另外，预设数量可以按照需要选择，以能够满足需要为准。

举一例子便于进行理解：假设预设数量为 100，则可以确定与当前处理的真实图像相似的 100 张历史图像，分别将 100 张历史图像除背景以外的内容擦除干净，并计算各自对应的信息熵的值，得到 100 个信息熵的值，从 100 个信息熵的值中确定最小的信息熵的值，并将之作为预设信息熵阈值。应理解，这个例子仅仅是便于理解的示例，并不作为对本申请的任何限制。

在一些可选的实施例中，若擦除后图像的信息熵不满足预设信息熵阈值，则调整形态学处理参数，以对融合掩膜进行形态学处理，并使用形态学处理完成后的融合掩膜，返回步骤 S108 中的“从真实图像中，擦除融合掩膜对应的图像区域”的步骤重新执行。

具体地，在擦除后图像的信息熵不满足预设信息熵阈值时，调整形态学处理参数，以对融合掩膜进行形态学处理，并使用形态学处理完成后的融合掩膜，使得返回从真实图像中，擦除融合掩膜对应的图像区域重新执行，从而便于直至将不满足预设信息熵阈值的擦除后图像的信息熵调整至满足预设信息熵阈值，以便于可以保证真实图像中的融合掩膜对应的图像区域擦除干净，将擦除干净得到的擦除后图像作为真实图像的背景图像使用，能够使得生成的扫描图像的背景效果更好，并且后续基于背景图像来合成扫描图像以作为训练样本时，也能够使得实现软件扫描的机器学习模型的训练效果更好。

图 6 为本申请实施例中的一种场景示例的示意图。再参照图 6 所示对本申请的背景图像生成方法进行总体理解，图 6 中示出了一个基于物理文件生成的真实图像（如图 6 中左侧界面所示），真实图像中包括文字类型内容、图标类型内容、色块图案类型内容和印章类型内容，经过本申请中的背景图像生成方法，输出得到的背景图像（如图 6 中右侧界面所示），其中擦除了上述文字类型内容、图标类型内容、色块图案类型内容和印章类型内容。应理解，图 6 中的示例并不作为对本申请中的任何限制。

在一些可选的实施例中，所述背景图像生成方法还包括：基于背景图像和预设的文件内容图像，合成扫描文件样本图像。本申请通过这种方式，可以有效地合成扫描文件

样本图像，以便于进行使用。

具体地，预设的文件内容图像可以是已有的能够表示文件内容的图像。预设的文件内容图像也可以包括从图像（这里所说的图像可以包括但不限于真实图像）中分割出的表示文件内容的图像。例如，文件内容可以包括文字内容、图表内容、色块图案内容、印章内容等中的至少之一。预设的文件内容图像在被分割出来后可以存储在合适的存储空间中，以便于在需要的时候获取。可选地，预设的文件内容图像可以是纯色背景或者透明背景，这样能更方便与步骤 S108 中生成的背景图像进行合成，得到扫描文件样本图像。在一个可选方式中，上述纯色背景可以为白色背景，白色背景与实际的扫描文件背景更为贴近，也更易于合成处理。

10 参照图 7A 所示的场景示例进行理解，其示出了基于背景图像和预设的文件内容图像，合成扫描文件样本图像的一个示例性过程，应理解，图 7A 中所示的示例并不作为对本申请实施例中的任何限制。

可选地，可以通过将预设的文件内容图像覆盖或者粘贴到背景图像上的合适区域，从而合成得到扫描文件样本图像。

15 或者，在另一些可选的实施例中，所述背景图像生成方法还包括：基于背景图像和预设的文件内容文本，合成扫描文件样本图像。本申请通过这种方式，也同样可以有效地合成扫描文件样本图像，以便于进行使用。

具体地，预设的文件内容文本可以是按照需要生成的文本内容。需要说明的是，文件内容文本与上述文件内容图像区别在于，文件内容图像是已有的图像，而文件内容文本是按照需要进行生成。举一例子来说，如果用户希望扫描文件样本图像的正中间区域包括“发票”二字，则用户可以通过软件输入“发票”二字的文本内容，并调整合适的文本格式（例如字体、字号等），并将这样的“发票”文本内容通过手动或者自动的方式调整到背景图像的正中间区域，从而实现基于背景图像和预设的文件内容文本，合成满足需要的扫描文件样本图像；应理解，这个示例并不作为对本申请中的任何限制。

25 参照图 7B 所示的场景示例进行理解，其示出了基于背景图像和预设的文件内容文本，合成扫描文件样本图像的一个示例性过程，应理解，图 7B 中所示的示例并不作为对本申请实施例中的任何限制。

在一些可选的实施例中，所述背景图像生成方法还包括：基于扫描文件样本图像，对用于生成扫描文件的机器学习模型进行训练。

30 以本申请技术方案得到的扫描文件样本图像作为训练样本，训练样本的背景图像基于真实图像生成，因而在数据分布上和真实的样本分布一致，因此训练样本的背景效果较好，并且能够使得对用于生成扫描文件的机器学习模型的训练效果更好，另外，由于通过本申请中的技术方案可以高效生成背景图像及扫描文件样本图像（即训练样本），因此也方便于快速地且针对性地帮助机器学习模型进行迭代。

35 综上所述，本申请实施例中提供的方案，可以基于真实的包含文件信息的图像，对其进行目标内容检测和掩膜处理，并基于多个目标内容掩膜生成融合掩膜，基于融合掩

膜进行图像区域擦除，从而获得真实图像对应的背景图像。由此，一方面，无需图像处理人员进行复杂的参数学习和调试，基于真实图像即可生成相应的背景图像，降低了背景图像生成成本；另一方面，背景图像基于真实图像生成，因而在数据分布上和真实的样本分布一致，不仅能够使得生成的扫描图像的背景效果较好；并且，后续，基于背景图像来合成扫描图像以作为训练样本时，也能够使得实现软件扫描的机器学习模型的训练效果更好。

可以理解的是，前述对背景图像生成方法进行的说明，仅作为对本申请的一些示例性的说明，并不作为对本申请的任何限制。

根据本申请实施例的第二方面，提供了一种电子设备，包括：处理器、存储器、通信接口和通信总线，所述处理器、所述存储器和所述通信接口通过所述通信总线完成相互间的通信；所述存储器用于存放至少一可执行指令，所述可执行指令使所述处理器执行如第一方面所述方法对应的操作。参照图 8，示出了根据本申请实施例的一种电子设备的结构示意图，本申请具体实施例并不对电子设备的具体实现做限定。

如图 8 所示，该电子设备 800 可以包括：处理器(processor)802、通信接口 (Communications Interface)804、存储器(memory)806、以及通信总线 808。其中：

处理器 802、通信接口 804、以及存储器 806 通过通信总线 808 完成相互间的通信。

通信接口 804，用于与其它电子设备或服务器进行通信。

处理器 802，用于执行程序 810，具体可以执行上述背景图像生成方法实施例中的相关步骤。

具体地，程序 810 可以包括程序代码，该程序代码包括计算机操作指令。

处理器 802 可能是 CPU，或者是 GPU（Graphic Processing Unit，图形处理器）或者是特定集成电路 ASIC（Application Specific Integrated Circuit），或者是被配置成实施本申请实施例的一个或多个集成电路。智能设备包括的一个或多个处理器，可以是同一类型的处理器，如一个或多个 CPU；也可以是不同类型的处理器，如一个或多个 CPU 以及一个或多个 ASIC。

存储器 806，用于存放程序 810。存储器 806 可能包含高速 RAM 存储器，也可能还包括非易失性存储器（non-volatile memory），例如至少一个磁盘存储器。

程序 810 可包括多条计算机指令，程序 810 具体可以通过多条计算机指令使得处理器 802 执行前述多个方法实施例中任一实施例所描述的背景图像生成方法对应的操作。

程序 810 中各步骤的具体实现可以参见上述方法实施例中的相应步骤和单元中对应的描述，并具有相应的有益效果，在此不赘述。所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的设备和模块的具体工作过程，可以参考前述方法实施例中的对应过程描述，在此不再赘述。

根据本申请实施例的第三方面，本申请实施例还提供了一种计算机存储介质，其上存储有计算机程序，该程序被处理器执行时实现前述多个方法实施例中任一实施例所描述的方法。该计算机存储介质包括但不限于：只读光盘（Compact Disc Read-Only Memory，

CD-ROM)、随机存储器(Random Access Memory, RAM)、软盘、硬盘或磁光盘等。

本申请实施例还提供了一种计算机程序产品,包括计算机指令,该计算机指令指示计算设备执行上述多个方法实施例中的任一背景图像生成方法对应的操作。

本申请实施例中的电子设备 800/计算机存储介质/计算机程序产品实施例,在前述的
5 背景图像生成方法实施例中已经详细说明,因此其相关内容及有益效果可参照上述的方法实施例进行理解,在此不再进行赘述。

此外,需要说明的是,本申请实施例所涉及到的与用户有关的信息(包括但不限于用户设备信息、用户个人信息等)和数据(包括但不限于用于对模型进行训练的样本数据、用于分析的数据、存储的数据、展示的数据等),均为经用户授权或者经过各方充分授权的信息和数据,并且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准,并提供有相应的操作入口,供用户选择授权或者拒绝。
10

需要指出,根据实施的需要,可将本申请实施例中描述的各个部件/步骤拆分为更多部件/步骤,也可将两个或多个部件/步骤或者部件/步骤的部分操作组合成新的部件/步骤,以实现本申请实施例的目的。

上述根据本申请实施例的方法可在硬件、固件中实现,或者被实现为可存储在记录介质(诸如 CD-ROM、RAM、软盘、硬盘或磁光盘)中的软件或计算机代码,或者被实现通过网络下载的原始存储在远程记录介质或非暂时机器可读介质中并将被存储在本地记录介质中的计算机代码,从而在此描述的方法可被存储在使用通用计算机、专用处理器或者可编程或专用硬件(诸如专用集成电路(Application Specific Integrated Circuit, ASIC)或现场可编程门阵列(Field Programmable Gate Array, FPGA))的记录介质上的这样的软件处理。可以理解,计算机、处理器、微处理器控制器或可编程硬件包括可存储或接收软件或计算机代码的存储组件(例如,随机存储器(Random Access Memory, RAM)、只读存储器(Read-Only Memory, ROM)、闪存等),当所述软件或计算机代码被计算机、处理器或硬件访问且执行时,实现在此描述的方法。此外,当通用计算机访问用于实现在此示出的方法的代码时,代码的执行将通用计算机转换为用于执行在此示出的方法的专用计算机。
15
20
25

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方法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能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特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本申请实施例的范围。
30

以上实施方式仅用于说明本申请实施例,而并非对本申请实施例的限制,有关技术领域的普通技术人员,在不脱离本申请实施例的精神和范围的情况下,还可以做出各种变化和变型,因此所有等同的技术方案也属于本申请实施例的范畴,本申请实施例的专利保护范围应由权利要求限定。
35

权利要求书

- 1.一种背景图像生成方法，包括：
获取基于物理文件生成的真实图像；
对所述真实图像进行多种不同类型的目标内容检测，获得对应的多个目标内容图像
5 区域及对应的多个目标内容掩膜；
对所述多个目标内容掩膜进行融合，获得融合掩膜；
从所述真实图像中，擦除所述融合掩膜对应的图像区域，并根据擦除结果获得所述
真实图像对应的背景图像。
- 2.根据权利要求 1 所述的方法，其中，所述对所述多个目标内容掩膜进行融合，获
10 得融合掩膜，包括：
根据所述多个目标内容掩膜中各目标内容掩膜的置信度，对所述多个目标内容掩膜
进行过滤；
对过滤后的目标内容掩膜进行融合，获得融合掩膜。
- 3.根据权利要求 2 所述的方法，其中，
15 在所述获得融合掩膜之后，所述方法还包括：根据形态学处理参数，对所述融合掩
膜进行形态学处理；
所述从所述真实图像中，擦除所述融合掩膜对应的图像区域，包括：从所述真实图
像中，擦除进行了形态学处理的融合掩膜对应的图像区域。
- 4.根据权利要求 1-3 任一项所述的方法，其中，所述从所述真实图像中，擦除所述融
20 合掩膜对应的图像区域，并根据擦除结果获得所述真实图像对应的背景图像，包括：
从所述真实图像中，擦除所述融合掩膜对应的图像区域，获得擦除后图像；
计算所述擦除后图像的信息熵，若所述信息熵满足预设信息熵阈值，则将所述擦除
后图像作为所述真实图像对应的背景图像。
- 5.根据权利要求 4 所述的方法，其中，所述方法还包括：
25 若所述信息熵不满足所述预设信息熵阈值，则调整形态学处理参数，以对所述融合
掩膜进行形态学处理，并使用形态学处理完成后的融合掩膜，返回所述从所述真实图像
中，擦除所述融合掩膜对应的图像区域的步骤重新执行。
- 6.根据权利要求 4 所述的方法，其中，所述预设信息熵阈值通过以下方式确定：
确定与当前处理的真实图像相似的预设数量的历史图像，根据所述历史图像对应的
30 背景图像的信息熵中的最小值，确定所述预设信息熵阈值。
- 7.根据权利要求 1 所述的方法，其中，所述目标内容包括以下类型至少之一：文字
类型内容、图表类型内容、色块图案类型内容、印章类型内容；
所述对所述真实图像进行多种不同类型的目标内容检测，获得对应的多个目标内容
图像区域及对应的多个目标内容掩膜，包括：
35 针对所述真实图像进行以下目标内容检测至少之二：针对文字的检测、针对图表的
检测、针对色块图案的检测、针对印章的检测；

- 根据检测结果，获得检测出的多种类型的目标内容对应的多个图像区域；
对所述多个图像区域进行掩膜，获得对应的多个目标内容掩膜。
- 8.根据权利要求 7 所述的方法，其中，所述根据检测结果，获得检测出的多种类型的目标内容对应的多个图像区域，包括以下至少之一：
- 5 若根据检测结果，确定检测出文字类型内容，则对所述文字类型内容进行文字识别，以根据文字识别结果，确定各个文字分别对应的图像区域；
 若根据检测结果，确定检测出图表类型内容，则对所述图表类型内容进行图像分割，以获得所述图表类型内容对应的图像区域；
 若根据检测结果，确定检测出色块图案类型内容，则对所述色块图案类型内容进行
10 图像分割，以获得所述色块图案内容对应的图像区域；
 若根据检测结果，确定检测出印章类型内容，则对所述印章类型内容进行文字识别及边框识别，以根据识别结果，确定印章对应的图像区域。
- 9.根据权利要求 1-3 任一项所述的方法，其中，所述方法还包括：
 基于所述背景图像和预设的文件内容图像，合成扫描文件样本图像；
15 或者，
 基于所述背景图像和预设的文件内容文本，合成扫描文件样本图像。
- 10.根据权利要求 9 所述的方法，其中，所述方法还包括：
 基于所述扫描文件样本图像，对用于生成扫描文件的机器学习模型进行训练。
- 11.一种电子设备，包括：处理器、存储器、通信接口和通信总线，所述处理器、所
20 述存储器和所述通信接口通过所述通信总线完成相互间的通信；
 所述存储器用于存放至少一可执行指令，所述可执行指令使所述处理器执行如权利
 要求 1-10 中任一项所述的方法对应的操作。
- 12.一种计算机存储介质，其上存储有计算机程序，该程序被处理器执行时实现如权利
 要求 1-10 中任一所述的方法。
- 25 13.一种计算机程序产品，包括计算机指令，该计算机指令指示计算设备执行如权利
 要求 1-10 中任一所述的方法对应的操作。