

5

图像处理模型的训练方法、图像处理方法及装置

本申请要求于 2023 年 12 月 27 日提交中国专利局、申请号为 202311813675.3、发明名称为“图像处理模型的训练方法、图像处理方法及装置”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

本申请涉及图像处理领域，更在一种可能的实现中涉及一种图像处理模型的训练方法、图像处理方法、图像处理模型的训练装置、图像处理装置、电子设备和存储介质。

10

背景技术

随着科技的发展，图像处理技术越来越多地应用于各个领域。因为图像中的文本承载更明晰的信息，所以关于图像中的文本处理一直是大家关注的热点。相关技术中，一些人工智能模型已经应用于图像中的文本处理，其具有准确度高、速度快等优点。

15

但因为文本具有字体、字形、大小等各种变化，通常需要针对特定的应用场景训练专用的人工智能模型。在模型的训练时，常规方法是依靠人工对收集的大量图像数据进行标注，然后基于标注结果进行训练。上述人工标注的过程耗时费力，导致训练周期也较长。

20

发明内容

考虑到上述问题而提出了本申请。

25

根据本申请一个方面，提供了一种图像处理模型的训练方法，图像处理模型包括文本检测模型和文本判断模型，文本检测模型用于对图像进行文本识别，文本判断模型用于检测图像中的文本的真实性，所述训练方法包括：

步骤 S110，获取第一图像和对应的正样本标签，其中，第一图像中包括第一文本，正样本标签包括第一文本的真实性的信息；

步骤 S120，利用文本生成工具，获得第二图像以及对应的负样本标签，其中，第二图像中包括利用文本生成工具所生成的第二文本，负样本标签包括

第二文本的真实性的信息和文本识别信息；

步骤 S130，将第一图像和第二图像均输入到文本判断模型，以由文本判断模型输出检测结果，将第二图像输入到文本检测模型，以由文本检测模型输出文本识别结果，基于正样本标签、负样本标签、文本识别结果和检测结果，
5 计算图像处理模型的损失值，利用损失值调整图像处理模型对应的参数，以训练图像处理模型。

在一种可能的实现方式中，损失值包括文本判断模型的第一损失值和文本检测模型的第二损失值，

步骤 S130 包括：首先，将第一图像和第二图像均输入到文本判断模型，
10 以由文本判断模型输出检测结果，基于正样本标签、负样本标签和检测结果，计算文本判断模型的第一损失值，利用第一损失值调整文本判断模型对应的参数，以训练文本判断模型；

然后，将第二图像输入到文本检测模型，以由文本检测模型输出文本识别结果，基于负样本标签和文本识别结果，计算文本检测模型的第二损失值，利用第二损失值调整文本检测模型对应的参数，以训练文本检测模型。
15

在一种可能的实现方式中，利用损失值调整图像处理模型对应的参数，以训练图像处理模型，包括：利用损失值同时调整文本判断模型对应的参数和文本检测模型对应的参数，以训练图像处理模型。

在一种可能的实现方式中，基于正样本标签、负样本标签、文本识别结果和检测结果，计算图像处理模型的损失值，包括：基于正样本标签、负样本标签和检测结果，计算文本判断模型的第一损失值；基于负样本标签和文本识别结果，计算文本检测模型的第二损失值；基于第一损失值和第二损失值，计算图像处理模型的损失值。
20

在一种可能的实现方式中，基于第一损失值和第二损失值，计算图像处理模型的损失值，包括：对第一损失值和第二损失值进行加权求和，以将所计算的和确定为图像处理模型的损失值。
25

在一种可能的实现方式中，文本检测模型和文本判断模型具有共享权重网络。

在一种可能的实现方式中，获得第二图像以及对应的负样本标签，包括：

获取背景图像；利用文本生成工具，生成第二文本并获得第二文本对应的负样本标签；将第二文本映射到背景图像，以生成第二图像。

根据本申请的又一个方面，提供了一种图像处理方法，处理方法包括：获取待处理图像，待处理图像中包括文本；将待处理图像输入上述训练方法所训练的图像处理模型，以输出待处理图像的文本识别结果和/或待处理图像中的文本的真实性检测结果。

根据本申请的又一个方面，提供了一种图像处理模型的训练装置，图像处理模型包括文本检测模型和文本判断模型，文本检测模型用于对图像进行文本识别，文本判断模型用于检测图像中的文本的真实性；训练装置包括：

第一获取模块，用于获取第一图像和对应的正样本标签，其中，第一图像中包括第一文本，正样本标签包括第一文本的真实性的信息；

第二获取模块，用于利用文本生成工具，获得第二图像以及对应的负样本标签，其中，第二图像中包括利用文本生成工具所生成的第二文本，所述负样本标签包括所述第二文本的真实性的信息和文本识别信息；

训练模块，用于将所述第一图像和所述第二图像均输入到所述文本判断模型，以由文本判断模型输出检测结果，将第二图像输入到文本检测模型，以由文本检测模型输出文本识别结果，基于正样本标签、负样本标签、文本识别结果和检测结果，计算图像处理模型的损失值，利用损失值调整图像处理模型对应的参数，以训练图像处理模型。

根据本申请的又一个方面，提供了一种图像处理装置，处理装置包括：第三获取模块，用于获取待处理图像，待处理图像中包括文本；处理模块，用于将待处理图像输入上述训练方法所训练的图像处理模型，以输出待处理图像的文本识别结果和/或待处理图像中的文本的真实性检测结果。

根据本申请的又一个方面，提供了一种电子设备，包括处理器和存储器，存储器中存储有计算机程序指令，计算机程序指令被处理器运行时用于执行上述图像处理模型的训练方法和/或图像处理方法。

根据本申请的再一个方面，提供了一种存储介质，在存储介质上存储了程序指令，程序指令在运行时用于执行上述图像处理模型的训练方法和/或图像处理方法。

上述训练图像处理模型的过程中，不仅利用了包括人工字符的第二图像，也利用了包括真实字符的第一图像。如果仅利用第二图像进行模型训练，其对真实图像虽然具备一定的识别能力，但是识别精度较低。在上述实施例中，利用文本生成工具获得第二图像以及对应的负样本标签，基于第一图像和其对应的正样本标签以及第二图像以及对应的负样本标签训练图像处理模型。由此，基于文本生成工具所生成的语义标签来进行图像处理模型的训练。该方案有效地节省了工作人员的精力，加速了模型训练的速度。

上述说明仅是本申请技术方案的概述，为了能够更清楚了解本申请的技术手段，而可依照说明书的内容予以实施，并且为了让本申请的上述和其它目的、特征和优点能够更明显易懂，以下特举本申请的具体实施方式。

附图说明

通过结合附图对本申请实施例进行更详细的描述，本申请的上述以及其它目的、特征和优势将变得更加明显。附图用来提供对本申请实施例的进一步理解，并且构成说明书的一部分，与本申请实施例一起用于解释本申请，并不构成对本申请的限制。在附图中，相同的参考标号通常代表相同部件或步骤。

图 1 示出了根据本申请一个实施例的图像处理模型的训练方法的示意性流程图；

图 2 示出了根据本申请一个实施例的第二图像；

图 3 示出了根据本申请一个实施例的利用图像处理模型进行图像处理的示意图；

图 4 示出了根据本申请另一个实施例的第二图像；

图 5 示出了根据本申请一个实施例的图像处理模型的训练装置的示意性框图；

图 6 示出了根据本申请一个实施例的图像处理装置的示意性框图；

图 7 示出了根据本申请一个实施例的电子设备的示意性框图。

具体实施方式

为了使得本申请的目的、技术方案和优点更为明显，下面将参照附图详细

描述根据本申请的示例实施例。显然，所描述的实施例仅仅是本申请的一部分实施例，而不是本申请的全部实施例，应理解，本申请不受这里描述的示例实施例的限制。基于本申请中描述的本申请实施例，本领域技术人员在没有付出创造性劳动的情况下所得到的所有其它实施例都应落入本申请的保护范围之内。

为了至少部分地解决上述技术问题，根据本申请的一个方面，提供了一种图像处理模型的训练方法。该图像处理模型用于对包括文本的图像进行处理。图像处理模型包括文本识别模型文本检测模型和文本检测模型文本判断模型。

文本识别模型文本检测模型用于对图像进行文本识别。在一种可能的实现中，文本检测模型可以对图像进行文本定位并识别所定位的文本。文本检测模型可以包括卷积循环神经网络（CRNN）和链接时序分类模型。

文本判断模型用于检测图像中的文本的真实性。文本判断模型可以用来区分输入的图像是否为合成的假图像。文本判断模型可以包括指针生成网络（PGNet）、深度双边网络（DBNet）、基于字符区域感知的文本检测（CRAFT）等网络模型，也可以是全卷积单阶段目标检测（FCOS）、实时目标检测（YOLO）等目标检测网络模型。文本判断模型能够辅助文本检测模型进行更准确的文本识别。

图 1 示出了根据本申请一个实施例的图像处理模型的训练方法的示意性流程图。如图 1 所示，该方法包括步骤 S110、步骤 S120 和步骤 S130。

步骤 S110，获取第一图像和对应的正样本标签。第一图像中包括第一文本。正样本标签包括第一文本的真实性的信息。

第一图像可以为原始采集的或者原始采集的图像经处理获得的、包括工业领域中有限的字符集的图像，例如，带有生产日期的图像或者带有产品编号的图像等。第一图像可以是 RGB 图像或灰度图像。第一图像可以是图像采集装置直接采集到的原始图像，也可以是对原始图像进行预处理操作后的图像。该预处理操作可以包括为了改善图像的视觉效果，提高其清晰度，或是突出图像中的文本的所有操作。示例性而非限制性地，预处理操作可以包括对原始图像的数字化的、几何变换、归一化、滤波等操作。第一图像中包括第一文本。因为第一图像是原始采集的或者原始采集的图像经处理获得的，所以其中的第一文

本包括真实字符。该真实字符可以包括数字和字母等。

第一图像具有相对应的正样本标签，正样本标签可以指示第一文本的真实性。如前所述，因为第一图像均为真实图像，所以第一文本均为真实文本。

步骤 S120，利用文本生成工具，获得第二图像以及对应的负样本标签，
5 其中，第二图像中包括利用文本生成工具所生成的第二文本。负样本标签包括第二文本的真实性的信息和文本识别信息。

在本申请的实施例中，可以在工业场景中采集少量的干净背景的图像，其中不包括任何文本。然后，可以利用文本生成工具在图像中的指定位置自动生成各种类型的文本内容。图 2 示出了根据本申请一个实施例的第二图像。图 2
10 所示第二图像中的文本为经文本生成工具生成的字符。为方便描述，在下文中称利用文本生成工具所生成的第二文本中的字符为人工字符。文本生成工具可以包括但不限于 TTF 文本生成工具。在本申请的实施例中，对文本生成工具的种类不做任何限定。凡是自动化生成文本内容的工具均在本申请保护范围之内。

15 第二图像中的人工字符可以是与第一图像中的第一文本中的真实字符相对应的。例如，第二图像中的人工字符具有与第一图像中的真实字符相同的字体、形状或大小等样式。由此，有助于文本判断模型和文本检测模型分别进行文本检测和文本识别。

第二图像也具有其对应的负样本标签。负样本标签可以指示第二文本的
20 真实性的信息以及文本识别信息。其中，因为第二图像是利用文本生成工具生成的，所以其中的第二文本可以认为是假的，即第二图像也是假的。文本识别信息可以包括图像中的文本位置信息、文本内容信息等。可以理解，文本生成工具在第二图像中生成人工字符的同时，即可同时生成对应的负样本标签。由此，避免了人工收集大量包括文本的图像以及对该图像进行标注的过程。

25 步骤 S130，将第一图像和第二图像均输入到文本判断模型，以由文本判断模型输出检测结果，将第二图像输入到文本检测模型，以由文本检测模型输出文本识别结果，基于正样本标签、负样本标签、文本识别结果和检测结果，计算图像处理模型的损失值，利用损失值调整图像处理模型对应的参数，以训练图像处理模型。

模型的损失值可以用来衡量模型预测结果与真实标签之间的差异程度。损失值越小，表示模型的预测结果与真实标签越接近，也就意味着模型的性能越好。当模型的损失值较大时，可以调整模型的参数，从而提高模型的准确性和泛化能力。

- 5 文本判断模型输出的检测结果可以表示输入图像为真图像或者假图像。其中，真图像为其中包括真实字符的图像，假图像为其中包括人工字符的图像。第一图像相对应的正样本标签可以表示第一图像为真图像，第二图像相对应的负样本标签可以表示第二图像为假图像。因此可以根据正样本标签、负样本标签以及文本判断模型输出的检测结果计算文本判断模型的损失函数的损失值。
- 10 可以基于该损失值对文本判断模型的参数进行调整，直至文本判断模型输出的检测结果较大概率地与输入图像的标签相一致，即文本判断模型能够准确地检测出图像中的文本的真实性为止。

- 文本检测模型输出图像的文本识别结果，具体可以包括文本在图像中的位置以及文本的具体内容等。第二图像相对应的负样本标签可以表示第二图像中人工字符的在第二图像中的位置和具体内容。可以根据负样本标签以及文本检测模型输出的文本识别结果计算文本检测模型的损失函数的损失值。可以基于该损失值对文本检测模型的参数进行调整，直至文本检测模型输出的文本识别结果较大概率地与第二图像的负样本标签相一致，即文本检测模型能够准确地识别出图像中的文本为止。
- 15

- 20 训练好的文本判断模型和训练好的文本检测模型构成了训练好的图像处理模型。

- 上述训练图像处理模型的过程中，不仅利用了包括人工字符的第二图像，也利用了包括真实字符的第一图像。如果仅利用第二图像进行模型训练，其对真实图像虽然具备一定的识别能力，但是识别精度较低。在上述实施例中，利用文本生成工具获得第二图像以及对应的负样本标签，基于第一图像和其对应的正样本标签以及第二图像以及对应的负样本标签训练图像处理模型。由此，基于文本生成工具所生成的语义标签来进行图像处理模型的训练。该方案有效地节省了工作人员的精力，加速了模型训练的速度。
- 25

 在一些实施例中，图像处理模型中的文本检测模型和文本判断模型具有共

享权重网络。

文本检测模型和文本判断模型都是神经网络模型。神经网络是由大量的神经元相互连接形成的。每个神经元在接收到输入后，会进行线性加权处理。这些线性加权处理可以通过权重来表示。其中，当一个神经元有两个输入时，每个输入都会乘以一个关联的权重，然后加在一起作为输出的结果。这些权重在神经网络训练过程中可以通过随机初始化，并在模型训练过程中进行更新。

在文本检测模型和文本判断模型中包括共享权重的网络。换言之，这部分共享权重网络之间的权重保持一致。在训练过程中，共享权重网络中的一个网络的神经元的权重改变，则另一个网络的对应神经元的权重也相应地改变。例如，该共享权重网络可以称为骨干网络，可以用于提取图像的特征。

图 3 示出了根据本申请一个实施例的利用图像处理模型进行图像处理的示意图。在图 3 中，第一共享权重网络和第一子网络构成文本检测模型。可以理解，第一子网络可以是文本检测模型的下游任务网络。第二共享权重网络和第二子网络构成文本判断模型。因为文本判断模型用于检测图像中的文本的真实性，所以第二子网络可以利用判别器实现。如图 3 所示，将第二图像和第一图像分别输入到第一共享权重网络和第二共享权重网络。对于经第一共享权重网络计算的第二图像，将输入到第一子网络和第二子网络二者。而对于经第二共享权重网络计算的第一图像，将仅输入到第二子网络。从文本判断模型角度来说，其接收了第一图像和第二图像，分别输出了二者的真实性。从文本检测模型角度来说，其仅接收了第二图像，输出了第二图像中的文本识别结果。

在一种可能的实现方式中，文本识别结果可以包括图像中的文本的位置信息、文本中每个字符的位置信息以及文本内容信息。可以理解文本的位置信息可以利用文本的外接矩形框来表示。为方便描述，以下称该矩形框为文本框。在一种可能的实现中，可以利用该文本框的对顶角的顶点的坐标来表示，例如左上顶点坐标和右下顶点坐标。替代地，该文本框也可以利用其中心点以及其长和宽来表示。字符的位置信息可以利用字符的中心位置的坐标来表示。文本内容信息表示了文本可以包括哪些字符，例如诸如 1、2、3 等的数字字符和诸如 a、b、c 等的英文字符等。基于上述图像处理模型的输出结果以及上述正样本标签和负样本标签，能够对图像处理模型进行训练。

在上述技术方案中，图像处理模型不仅包括文本检测模型还包括文本判断模型，而且两个模型还具有共享权重网络。这使得共享权重网络不仅提取到第一图像的特征，还提取到第二图像的特征。由此，保证了图像处理模型的图像处理准确度。此外，共享权重网络可以有效地减少图像处理模型所需参数的数量，提高了模型的泛化能力，并且能够将从一个训练任务中学习到的知识应用到另一个相关任务中，进一步地提高了图像处理模型的训练效率。

在一些实施例中，步骤 S120 中获得第二图像以及对应的负样本标签，包括：获取背景图像；利用文本生成工具，生成第二文本并获得第二文本对应的负样本标签；将第二文本映射到背景图像，以生成第二图像。

背景图像可以是未印有文本的图像，第二文本可以在背景图像中的任意位置。利用文本生成工具，将生成的人工字符随机的映射到背景图像中，可以形成第二图像，同时还可以生成对应的负样本标签。第二图像可以包括较强的干扰信号，同时还具有类似真实场景下图像的特征。图 4 示出了根据本申请另一个实施例的第二图像。如图 4 所示，第二图像中的文本是利用文本生成工具生成人工字符后映射至背景图像中所生成的。该第二图像即具备较强的随机性，又具备真实图像所具备的特征。如此生成的第二图像有效地提高了图像处理模型的鲁棒性。

在上述实施例中，由于第二图像具有极强的随机性，并且自带标签信息，由此避免了人工收集大量带文本的图像和人工打标的操作，有效地提高了第二图像收集的速度，并且进一步提高了图像处理模型的训练效率。

在一些实施例中，损失值包括文本判断模型的第一损失值和文本检测模型的第二损失值。步骤 S130 包括以下步骤。首先，将第一图像和第二图像均输入到文本判断模型，以由文本判断模型输出检测结果，基于正样本标签、负样本标签和检测结果，计算文本判断模型的第一损失值，利用第一损失值调整文本判断模型对应的参数，以训练文本判断模型。然后，将第二图像输入到文本检测模型，以由文本检测模型输出文本识别结果，基于负样本标签和文本识别结果，计算文本检测模型的第二损失值，利用第二损失值调整文本检测模型对应的参数，以训练文本检测模型。

将第一图像和第二图像分别输入文本判断模型，以由文本判断模型输出各

自的检测结果。根据文本判断模型输出的检测结果、第一图像对应的正样本标签和第二图像对应的负样本标签计算文本判断模型的第一损失值。基于该第一损失值可以更新文本判断模型的参数。经过多次不断重复的计算、更新等步骤，以完成文本判断模型的训练。

- 5 在文本判断模型训练完成之后，仅将第二图像输入文本检测模型，以由文本检测模型输出文本识别结果。再根据文本检测模型输出的文本识别结果和负样本标签，计算文本检测模型的第二损失值，并利用第二损失值完成文本检测模型的训练。

- 10 在上述实施例中，首先对文本判断模型训练，而后再对文本检测模型进行训练，保证了二者的训练效果。特别地，对于文本判断模型和文本检测模型具有共享权重网络的实施例，在文本判断模型完成训练后，文本检测模型可以直接利用文本判断模型的部分参数进行训练，由此加快了图像处理模型的训练速度。

- 15 在一些实施例中，步骤 S130 中利用损失值调整图像处理模型对应的参数，以训练图像处理模型，包括：利用损失值同时调整文本判断模型对应的参数和文本检测模型对应的参数，以训练图像处理模型。

- 20 在上述实施例中，文本判断模型和文本检测模型可以在不同的计算单元上同时进行训练。换言之，将图像处理模型作为一个整体进行训练。同时训练文本判断模型和文本检测模型，可以有效地简化图像处理模型的训练流程，加快图像处理模型训练过程的速度，提升了训练效率。

- 25 在一些实施例中，步骤 S130 中基于正样本标签、负样本标签、文本识别结果和检测结果，计算图像处理模型的损失值，包括以下步骤 S131 至步骤 S133。在步骤 S131，基于正样本标签、负样本标签和检测结果，计算文本判断模型的第一损失值。在步骤 S132，基于负样本标签和文本识别结果，计算文本检测模型的第二损失值。在步骤 S133，基于第一损失值和第二损失值，计算图像处理模型的损失值。

第一损失值表示了检测结果与正样本标签和负样本标签之间的差异，由此，表示了文本判断模型的检测准确度。在一种可能的实现方式中，可以使用交叉熵损失函数计算上述第一损失值。

第二损失值表示了文本识别结果与负样本标签之间的差异，由此表示了文本检测模型的文本识别准确度。如前所述，文本识别结果可以包括文本框，该文本框用于表示图像中的文本的位置。替代地，该文本框也可以利用其中心点以及其长和宽来表示。第二损失值可以包括文本框分类损失、文本框回归损失和/或文本内容损失等。文本框分类损失主要用于判断文本框内是否包含文本。文本框回归损失用于优化文本框的位置和大小。对于每个文本框，需要计算其相对于真实文本框的偏移量。可以使用回归损失函数来计算文本框回归损失。文本内容损失用于确定文本检测模型对文本的识别能力。在一种可能的实现方式中，可以通过将文本框分类损失和文本框回归损失和文本内容损失进行加权求和，得到文本检测模型的第二损失值。文本检测模型的损失函数可以根据模型来确定。在一种可能的实现中，如采用 PGNet 作为文本检测模型，则可以采用 PGNet 中的损失函数。

在上述实施例中，通过分别计算第一损失值和第二损失值，并利用上述两个损失值计算图像处理模型的损失值，可以对图像处理模型的损失值进行精确的计算，有效地提高图像处理模型的稳定性和鲁棒性。

在一些实施例中，步骤 S133，基于第一损失值和第二损失值，计算图像处理模型的损失值，包括：对第一损失值和第二损失值进行加权求和，以将所计算的和确定为图像处理模型的损失值。

可以综合考虑第一损失值和第二损失值在图像处理模型中的相对重要性，来设置各自的权重。如果某个损失值对图像处理模型的性能有更大的贡献，可以赋予其更高的权重，从而更加重视该损失值。这有助于平衡不同损失值之间的关系，使得图像处理模型具有期望的表现。

在上述实施例中，根据第一损失值和第二损失值的权重进行加权求和，并将求和结果作为图像处理模型的损失值，用以训练图像处理模型，可以使图像处理模型具有较强的可靠性和鲁棒性。

根据本申请的再一方面，还提供了一种图像处理方法，处理方法包括以下步骤。首先，获取待处理图像，待处理图像中包括文本。然后，将待处理图像输入上述训练方法所训练的图像处理模型，以输出待处理图像的文本识别结果和/或待处理图像中的文本的真实性检测结果。

可以理解，该待处理图像可以是 RGB 图像或灰度图像。待处理图像可以是图像采集装置直接采集到的原始图像，也可以是对原始图像进行预处理操作后的图像。该预处理操作可以包括为了改善图像的视觉效果，提高其清晰度，或是突出图像中的文本的所有操作。示例性而非限制性地，预处理操作可以包括对原始图像的数字化、几何变换、归一化、滤波等操作。

在本申请的一些实施例中，利用上述图像处理方法，可以确定待处理图像中的文本识别结果。在该实施例中，图像处理模型中的文本判断模型的存在，有助于训练得到更准确的文本检测模型。由此，利用图像处理模型对待处理图像进行处理时，能够得到更准确的文本识别结果。

在本申请的一些实施例中，利用上述图像处理方法，可以判断待处理图像中的文本的真实性。在该实施例中，图像处理模型中的文本检测模型的存在，有助于训练得到更准确的文本判断模型。由此，利用图像处理模型对待处理图像进行处理时，能够得到更准确的文本判断结果。

在上述技术方案中，利用上述训练方法训练的图像处理模型输出待处理图像的文本识别结果和/或待处理图像中的文本的真实性检测结果。该方案保证了图像处理结果的准确性。

根据本申请的再一方面，还提供了一种图像处理模型的训练装置。该图像处理模型用于对包括文本的图像进行处理。图像处理模型包括文本检测模型和文本判断模型，文本检测模型用于对图像进行文本识别，文本判断模型用于检测图像中的文本的真实性。图 5 示出了根据本申请一个实施例的图像处理模型的训练装置的示意性框图。如图 5 所示，训练装置 500 包括第一获取模块 510、第二获取模块 520 和训练模块 530。

第一获取模块 510 用于获取第一图像和对应的正样本标签，第一图像中包括第一文本，正样本标签包括第一文本的真实性的信息。第二获取模块 520 用于利用文本生成工具，获得第二图像以及对应的负样本标签，其中，第二图像中包括利用所述文本生成工具所生成的第二文本，所述负样本标签包括所述第二文本的真实性的信息和文本识别信息。训练模块 530 用于将第一图像和第二图像均输入到文本判断模型，以由文本判断模型输出检测结果，将第二图像输入到文本检测模型，以由文本检测模型输出文本识别结果，基于正样本标签、

负样本标签、文本识别结果和检测结果，计算图像处理模型的损失值，利用损失值调整图像处理模型对应的参数，以训练图像处理模型。

根据本申请的再一方面，还提供了一种图像处理装置。图 6 示出了根据本申请一个实施例的图像处理装置的示意性框图。如图 6 所示，图像处理装置 5 600 包括第三获取模块 610 和处理模块 620。

第三获取模块 610 用于获取待处理图像，待处理图像中包括文本。处理模块 620 用于将待处理图像输入如上述训练方法所训练的图像处理模型，以输出待处理图像的文本识别结果和/或待处理图像中的文本的真实性检测结果。

根据本申请再一方面，还提供了一种电子设备 700。图 7 示出了根据本申请一个实施例的电子设备 700 的示意性框图。如图 7 所示，电子设备 700 包括 10 处理器 710 和存储器 720。其中，存储器 720 中存储有计算机程序指令，计算机程序指令被处理器运行时用于执行上述图像处理模型的训练方法和/或图像处理方法。

此外，根据本申请又一方面，还提供了一种存储介质，在存储介质上存储了程序指令，在程序指令被计算机或处理器运行时使得计算机或处理器执行 15 本申请实施例的上述图像处理模型的训练方法和/或图像处理方法的相应步骤，并且用于实现根据本申请实施例的上述图像处理模型的训练装置和/或图像处理装置的相应模块或上述电子设备中的相应模块。存储介质例如可以包括平板电脑的存储部件、个人计算机的硬盘、只读存储器(ROM)、可擦除可编程只读 20 存储器(EPROM)、便携式紧致盘只读存储器(CD-ROM)、USB 存储器、或者上述存储介质的任意组合。计算机可读存储介质可以是一个或多个计算机可读存储介质的任意组合。

本领域普通技术人员通过阅读上述图像处理模型的训练方法和/或图像处理方法的具体描述，能够理解上述图像处理模型的训练装置图像处理装置、电 25 子设备和存储介质的具体实现和有益效果，为了简洁，在此不再赘述。

实施例 1：一种图像处理模型的训练方法，所述图像处理模型包括文本检测模型和文本判断模型，所述文本检测模型用于对图像进行文本识别，所述文本判断模型用于检测图像中的文本的真实性；

所述训练方法包括：

步骤 S110, 获取第一图像和对应的正样本标签, 其中, 所述第一图像中包括第一文本, 所述正样本标签包括所述第一文本的真实性的信息;

步骤 S120, 利用文本生成工具, 获得第二图像以及对应的负样本标签, 其中, 所述第二图像中包括利用所述文本生成工具所生成的第二文本, 所述负样本标签包括所述第二文本的真实性的信息和文本识别信息;

步骤 S130, 将所述第一图像和所述第二图像均输入到所述文本判断模型, 以由所述文本判断模型输出检测结果, 将所述第二图像输入到文本检测模型, 以由所述文本识别模型输出文本识别结果, 基于所述正样本标签、所述负样本标签、所述文本识别结果和所述检测结果, 计算所述图像处理模型的损失值, 利用所述损失值调整所述图像处理模型对应的参数, 以训练所述图像处理模型。

实施例 2: 根据实施例 1 介绍的图像处理模型的训练方法, 所述损失值包括所述文本判断模型的第一损失值和所述文本检测模型的第二损失值,

所述步骤 S130 包括:

首先, 将所述第一图像和所述第二图像均输入到所述文本判断模型, 以由所述文本判断模型输出所述检测结果, 基于所述正样本标签、所述负样本标签和所述检测结果, 计算所述文本判断模型的第一损失值, 利用所述第一损失值调整所述文本判断模型对应的参数, 以训练所述文本判断模型;

然后, 将所述第二图像输入到所述文本检测模型, 以由所述文本检测模型输出所述文本识别结果, 基于所述负样本标签和所述文本识别结果, 计算所述文本检测模型的第二损失值, 利用所述第二损失值调整所述文本检测模型对应的参数, 以训练所述文本检测模型。

实施例 3: 根据实施例 1-2 中任一实施例介绍的图像处理模型的训练方法, 所述利用所述损失值调整所述图像处理模型对应的参数, 以训练所述图像处理模型, 包括:

利用所述损失值同时调整所述文本判断模型对应的参数和所述文本检测模型对应的参数, 以训练所述图像处理模型。

实施例 4: 根据实施例 1-3 中任一实施例介绍的图像处理模型的训练方法, 所述基于所述正样本标签、所述负样本标签、所述文本识别结果和所述检测结果, 计算所述图像处理模型的损失值, 包括:

基于所述正样本标签、所述负样本标签和所述检测结果，计算所述文本判断模型的第一损失值；

基于所述负样本标签和所述文本识别结果，计算所述文本检测模型的第二损失值；

5 基于所述第一损失值和所述第二损失值，计算所述图像处理模型的损失值。

实施例 5：根据实施例 1-4 中任一实施例介绍的图像处理模型的训练方法，所述基于所述第一损失值和所述第二损失值，计算所述图像处理模型的损失值，包括：

10 对所述第一损失值和所述第二损失值进行加权求和，以将所计算的和确定为所述图像处理模型的损失值。

实施例 6：根据实施例 1-5 中任一实施例介绍的图像处理模型的训练方法，所述文本检测模型和所述文本判断模型具有共享权重网络。

实施例 7：根据实施例 1-6 中任一实施例介绍的图像处理模型的训练方法，所述获得第二图像以及对应的负样本标签，包括：

15 获取背景图像；

利用文本生成工具，生成所述第二文本并获得所述第二文本对应的负样本标签；

将所述第二文本映射到所述背景图像，以生成所述第二图像。

实施例 8：一种图像处理方法，所述处理方法包括：

20 获取待处理图像，所述待处理图像中包括文本；

将所述待处理图像输入如实施例 1-7 任一实施例介绍的图像处理模型的训练方法所训练的图像处理模型，以输出所述待处理图像的文本识别结果和/或待处理图像中的文本的真实性检测结果。

25 实施例 9：一种图像处理模型的训练装置，所述图像处理模型包括文本检测模型和文本判断模型，所述文本检测模型用于对图像进行文本识别，所述文本判断模型用于检测图像中的文本的真实性；

所述训练装置包括：

第一获取模块，用于获取第一图像和对应的正样本标签，其中，所述第一图像中包括第一文本，所述正样本标签包括所述第一文本的真实性的信息；

第二获取模块，用于利用文本生成工具，获得第二图像以及对应的负样本标签，其中，所述第二图像中包括利用所述文本生成工具所生成的第二文本，所述负样本标签包括所述第二文本的真实性的信息和文本识别信息；

5 训练模块，用于将所述第一图像和所述第二图像均输入到所述文本判断模型，以由所述文本判断模型输出检测结果，将所述第二图像输入到文本检测模型，以由所述文本检测模型输出文本识别结果，基于所述正样本标签、所述负样本标签、所述文本识别结果和所述检测结果，计算所述图像处理模型的损失值，利用所述损失值调整所述图像处理模型对应的参数，以训练所述图像处理模型。

10 实施例 10：一种图像处理装置，所述处理装置包括：

第三获取模块，用于获取待处理图像，所述待处理图像中包括文本；

处理模块，用于将所述待处理图像输入如实施例 1-7 任一实施例介绍的图像处理模型的训练方法所训练的图像处理模型，以输出所述待处理图像的文本识别结果和/或待处理图像中的文本的真实性检测结果。

15 实施例 11：一种电子设备，包括处理器和存储器，所述存储器中存储有计算机程序指令，所述计算机程序指令被所述处理器运行时用于执行如实施例 1-7 任一实施例介绍的图像处理模型的训练方法和/或如实施例 8 所述的图像处理方法。

20 实施例 12：一种存储介质，在所述存储介质上存储了程序指令，所述程序指令在运行时用于执行如实施例 1-7 任一实施例介绍的图像处理模型的训练方法和/或如实施例 8 所述的图像处理方法。

25 尽管这里已经参考附图描述了示例实施例，应理解上述示例实施例仅仅是示例性的，并且不意图将本申请的范围限制于此。本领域普通技术人员可以在其中进行各种改变和修改，而不偏离本申请的范围和精神。所有这些改变和修改意在包括在所附权利要求所要求的本申请的范围之内。

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、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每个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的范围。

在本申请所提供的几个实施例中，应该理解到，所揭露的设备和方法，可以通过其它的方式实现。例如，以上所描述的设备实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个设备，或一些特征可以忽略，或不执行。

在此处所提供的说明书中，说明了大量具体细节。然而，能够理解，本申请的实施例可以在没有这些具体细节的情况下实践。在一些实例中，并未详细示出公知的方法、结构和技术，以便不模糊对本说明书的理解。

类似地，应当理解，为了精简本申请并帮助理解各个发明方面中的一个或多个，在对本申请的示例性实施例的描述中，本申请的各个特征有时被一起分组到单个实施例、图、或者对其的描述中。然而，并不应将该本申请的方法解释成反映如下意图：即所要求保护的本申请要求比在每个权利要求中所明确记载的特征更多的特征。更确切地说，如相应的权利要求书所反映的那样，其发明点在于可以用少于某个公开的单个实施例的所有特征的特征来解决相应的技术问题。因此，遵循具体实施方式的权利要求书由此明确地并入该具体实施方式，其中每个权利要求本身都作为本申请的单独实施例。

本领域的技术人员可以理解，除了特征之间相互排斥之外，可以采用任何组合对本说明书（包括伴随的权利要求、摘要和附图）中公开的所有特征以及如此公开的任何方法或者设备的所有过程或单元进行组合。除非另外明确陈述，本说明书（包括伴随的权利要求、摘要和附图）中公开的每个特征可以由提供相同、等同或相似目的的替代特征来代替。

此外，本领域的技术人员能够理解，尽管在此所述的一些实施例包括其它实施例中包括的某些特征而不是其它特征，但是不同实施例的特征的组合意味着处于本申请的范围之内并且形成不同的实施例。例如，在权利要求书中，所要求保护的实施例的任意之一都可以以任意的组合方式使用。

本申请的各个部件实施例可以以硬件实现，或者以在一个或者多个处理器上运行的软件模块实现，或者以它们的组合实现。本领域的技术人员应当理解，可以在实践中使用微处理器或者数字信号处理器（DSP）来实现根据本申请实

5 施例的图像处理模型的训练装置和图像处理装置中的一些模块的一些或者全部功能。本申请还可以实现为用于执行这里所描述的方法的一部分或者全部的装置程序（例如，计算机程序和计算机程序产品）。这样的实现本申请的程序可以存储在计算机可读介质上，或者可以具有一个或者多个信号的形式。这样的信号可以从因特网网站上下载得到，或者在载体信号上提供，或者以任何其他形式提供。

10 应该注意的是上述实施例对本申请进行说明而不是对本申请进行限制，并且本领域技术人员在不脱离所附权利要求的范围的情况下可设计出替换实施例。在权利要求中，不应将位于括号之间的任何参考符号构造成对权利要求的限制。单词“包含”不排除存在未列在权利要求中的元件或步骤。位于元件之前的单词“一”或“一个”不排除存在多个这样的元件。本申请可以借助于包括有若干不同元件的硬件以及借助于适当编程的计算机来实现。在列举了若干装置的单元权利要求中，这些装置中的若干个可以是通过同一个硬件项来具体体现。单词第一、第二、以及第三等的使用不表示任何顺序。可将这些单词解释为名
15 称。

以上所述，仅为本申请的具体实施方式或对具体实施方式的说明，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，可轻易想到变化或替换，都应涵盖在本申请的保护范围之内。本申请的保护范围应以权利要求的保护范围为准。

20

权利要求

1.一种图像处理模型的训练方法，其特征在于，所述图像处理模型包括文本检测模型和文本判断模型，所述文本检测模型用于对图像进行文本识别，所述文本判断模型用于检测图像中的文本的真实性；

5 所述训练方法包括：

 步骤 S110，获取第一图像和对应的正样本标签，其中，所述第一图像中包括第一文本，所述正样本标签包括所述第一文本的真实性的信息；

 步骤 S120，利用文本生成工具，获得第二图像以及对应的负样本标签，其中，所述第二图像中包括利用所述文本生成工具所生成的第二文本，所述负
10 样本标签包括所述第二文本的真实性的信息和文本识别信息；

 步骤 S130，将所述第一图像和所述第二图像均输入到所述文本判断模型，以由所述文本判断模型输出检测结果，将所述第二图像输入到文本检测模型，以由所述文本识别模型输出文本识别结果，基于所述正样本标签、所述负样本标签、所述文本识别结果和所述检测结果，计算所述图像处理模型的损失值，
15 利用所述损失值调整所述图像处理模型对应的参数，以训练所述图像处理模型。

2.根据权利要求 1 所述图像处理模型的训练方法，其特征在于，所述损失值包括所述文本判断模型的第一损失值和所述文本检测模型的第二损失值，

 所述步骤 S130 包括：

 首先，将所述第一图像和所述第二图像均输入到所述文本判断模型，以由
20 所述文本判断模型输出所述检测结果，基于所述正样本标签、所述负样本标签和所述检测结果，计算所述文本判断模型的第一损失值，利用所述第一损失值调整所述文本判断模型对应的参数，以训练所述文本判断模型；

 然后，将所述第二图像输入到所述文本检测模型，以由所述文本检测模型输出所述文本识别结果，基于所述负样本标签和所述文本识别结果，计算所述
25 文本检测模型的第二损失值，利用所述第二损失值调整所述文本检测模型对应的参数，以训练所述文本检测模型。

3.根据权利要求 1 所述图像处理模型的训练方法，其特征在于，所述利用所述损失值调整所述图像处理模型对应的参数，以训练所述图像处理模型，包括：

利用所述损失值同时调整所述文本判断模型对应的参数和所述文本检测模型对应的参数，以训练所述图像处理模型。

- 4.根据权利要求3所述图像处理模型的训练方法，其特征在于，
所述基于所述正样本标签、所述负样本标签、所述文本识别结果和所述检测结果，计算所述图像处理模型的损失值，包括：

基于所述正样本标签、所述负样本标签和所述检测结果，计算所述文本判断模型的第一损失值；

基于所述负样本标签和所述文本识别结果，计算所述文本检测模型的第二损失值；

- 10 基于所述第一损失值和所述第二损失值，计算所述图像处理模型的损失值。

5.根据权利要求4所述图像处理模型的训练方法，其特征在于，所述基于所述第一损失值和所述第二损失值，计算所述图像处理模型的损失值，包括：

对所述第一损失值和所述第二损失值进行加权求和，以将所计算的和确定为所述图像处理模型的损失值。

- 15 6.根据权利要求1至5任一项所述图像处理模型的训练方法，其特征在于，所述文本检测模型和所述文本判断模型具有共享权重网络。

7.根据权利要求1至5任一项所述图像处理模型的训练方法，其特征在于，所述获得第二图像以及对应的负样本标签，包括：

获取背景图像；

- 20 利用文本生成工具，生成所述第二文本并获得所述第二文本对应的负样本标签；

将所述第二文本映射到所述背景图像，以生成所述第二图像。

8.一种图像处理方法，其特征在于，所述处理方法包括：

获取待处理图像，所述待处理图像中包括文本；

- 25 将所述待处理图像输入如权利要求1至7任一项训练方法所训练的图像处理模型，以输出所述待处理图像的文本识别结果和/或待处理图像中的文本的真实性检测结果。

9.一种图像处理模型的训练装置，其特征在于，所述图像处理模型包括文本检测模型和文本判断模型，所述文本检测模型用于对图像进行文本识别，所

述文本判断模型用于检测图像中的文本的真实性；

所述训练装置包括：

第一获取模块，用于获取第一图像和对应的正样本标签，其中，所述第一图像中包括第一文本，所述正样本标签包括所述第一文本的真实性的信息；

- 5 第二获取模块，用于利用文本生成工具，获得第二图像以及对应的负样本标签，其中，所述第二图像中包括利用所述文本生成工具所生成的第二文本，所述负样本标签包括所述第二文本的真实性的信息和文本识别信息；

- 训练模块，用于将所述第一图像和所述第二图像均输入到所述文本判断模型，以由所述文本判断模型输出检测结果，将所述第二图像输入到文本检测模型，以由所述文本检测模型输出文本识别结果，基于所述正样本标签、所述负样本标签、所述文本识别结果和所述检测结果，计算所述图像处理模型的损失值，利用所述损失值调整所述图像处理模型对应的参数，以训练所述图像处理模型。
- 10

10.一种图像处理装置，其特征在于，所述处理装置包括：

- 15 第三获取模块，用于获取待处理图像，所述待处理图像中包括文本；
- 处理模块，用于将所述待处理图像输入如权利要求 1 至 7 任一项训练方法所训练的图像处理模型，以输出所述待处理图像的文本识别结果和/或待处理图像中的文本的真实性检测结果。

- 11.一种电子设备，包括处理器和存储器，其特征在于，所述存储器中存储有计算机程序指令，所述计算机程序指令被所述处理器运行时用于执行如权利要求 1 至 7 任一项所述的图像处理模型的训练方法和/或如权利要求 8 所述的图像处理方法。
- 20

- 12.一种存储介质，在所述存储介质上存储了程序指令，其特征在于，所述程序指令在运行时用于执行如权利要求 1 至 7 任一项所述的图像处理模型的训练方法和/或如权利要求 8 所述的图像处理方法。
- 25