



(12) 发明专利申请

(10) 申请公布号 CN 115035538 A

(43) 申请公布日 2022. 09. 09

(21) 申请号 202210685043.2 *G06V 30/18* (2022.01)

(22) 申请日 2022.03.22 *G06V 30/16* (2022.01)

(62) 分案原申请数据 *G06V 10/82* (2022.01)

202210279539.X 2022.03.22 *G06N 3/04* (2006.01)

G06N 3/08 (2006.01)

(71) 申请人 北京百度网讯科技有限公司 *G06F 40/30* (2020.01)

地址 100085 北京市海淀区上地十街10号
百度大厦2层

(72) 发明人 章成全 庾悦晨 李煜林 曹健健
钦夏孟 姚锬 韩钧宇 刘经拓
丁二锐 王井东

(74) 专利代理机构 北京同立钧成知识产权代理
有限公司 11205
专利代理师 马姣琴 臧建明

(51) Int. Cl.
G06V 30/19 (2022.01)

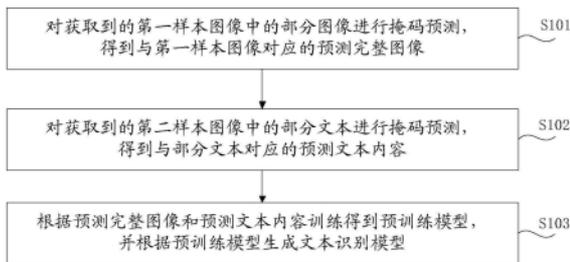
权利要求书4页 说明书14页 附图5页

(54) 发明名称

文本识别模型的训练方法、文本识别方法及装置

(57) 摘要

本公开提供了一种文本识别模型的训练方法、文本识别方法及装置,涉及人工智能技术领域,具体为深度学习、计算机视觉技术领域,可应用于光学字符识别等场景。方案为:对获取到的第一样本图像中的部分图像进行掩码预测,得到与第一样本图像对应的预测完整图像,对获取到的第二样本图像中的部分文本进行掩码预测,得到与部分文本对应的预测文本内容,根据预测完整图像和预测文本内容训练得到预训练模型,并根据预训练模型生成文本识别模型,文本识别模型用于对待识别图像进行文本识别,使得预训练模型学习到较强的图像视觉推理能力和文本语义推理能力,从而当基于预训练模型生成的文本识别模型进行文本识别时,提高文本识别的准确性和可靠性。



1. 一种文本识别模型的训练方法,其特征在于,包括:

对获取到的第一样本图像中的部分图像进行掩码预测,得到与所述第一样本图像对应的预测完整图像;

对获取到的第二样本图像中的部分文本进行所述掩码预测,得到与所述部分文本对应的预测文本内容,其中,所述第一样本图像和所述第二样本图像为不同的图像;

根据所述预测完整图像和所述预测文本内容训练得到预训练模型,并根据所述预训练模型生成文本识别模型,其中,所述文本识别模型用于对待识别图像进行文本识别。

2. 根据权利要求1所述的方法,其中,所述第一样本图像中的图像和所述第二样本图像中的文本所表征的内容不同。

3. 根据权利要求1所述的方法,其中,所述根据所述预训练模型生成文本识别模型,包括:

获取待识别任务和训练图像,其中,所述训练图像中包括文本;

根据所述待识别任务和所述训练图像,对所述预训练模型进行训练,得到所述文本识别模型。

4. 根据权利要求3所述的方法,其中,所述根据所述待识别任务和所述训练图像,对所述预训练模型进行训练,得到所述文本识别模型,包括:

将所述训练图像输入至所述预训练模型,得到所述训练图像对应的多模态特征图,其中,所述多模态特征图用于表征所述训练图像对应的图像特征和语义特征;

根据所述待识别任务和所述多模态特征图,生成所述文本识别模型。

5. 根据权利要求4所述的方法,其中,所述根据所述待识别任务和所述多模态特征图,生成所述文本识别模型,包括:

根据所述多模态特征图,预测所述训练图像在所述待识别任务下的预测识别结果;

根据所述训练图像预设的真实识别结果、以及所述预测识别结果,构建所述文本识别模型。

6. 根据权利要求1-5中任一项所述的方法,其中,所述掩码预测包括:

随机遮盖目标对象中的部分对象;

根据所述目标对象中未被遮盖的对象,对所述目标对象中被遮盖的部分对象进行预测,得到预测结果;

其中,若所述目标对象为第一样本图像,则所述目标对象中的部分对象为部分图像,所述预测结果为所述预测完整图像;若所述目标对象为第二样本图像,则所述目标对象中的部分对象为部分文本,所述预测结果为所述预测文本内容。

7. 根据权利要求6所述的方法,其中,根据所述目标对象中未被遮盖的对象,对所述目标对象中被遮盖的部分对象进行预测,得到预测结果,包括:

提取所述目标对象中未被遮盖的对象对应的对象特征,得到第一对象特征;

根据所述第一对象特征,对所述目标对象中被遮盖的部分对象进行预测,得到所述预测结果;

其中,若所述目标对象为第一样本图像,则所述第一对象特征为第一视觉特征;若所述目标对象为第二样本图像,则所述第一对象特征为第一语义特征。

8. 根据权利要求7所述的方法,其中,所述目标对象为第一样本图像,所述第一对象特

征为第一视觉特征；所述根据所述第一对象特征，对所述目标对象中被遮盖的部分对象进行预测，得到所述预测结果，包括

根据所述第一视觉特征，预测所述第一样本图像中被遮盖的部分图像对应的视觉特征，得到第二视觉特征；

根据所述第二视觉特征，确定所述第一样本图像中被遮盖的部分图像；

根据所述第一样本图像中未被遮盖的图像、以及确定出的所述第一样本图像中被遮盖的部分图像，生成所述预测完整图像。

9. 根据权利要求7或8所述的方法，其中，所述目标对象为第二样本图像，所述第一对象特征为第一语义特征；所述根据所述第一对象特征，对所述目标对象中被遮盖的部分对象进行预测，得到所述预测结果，包括：

根据所述第一语义特征，预测所述第二样本图像中被遮盖的部分文本对应的语义特征，得到第二语义特征；

根据所述第二语义特征，生成所述预测文本内容。

10. 一种文本识别方法，其特征在于，包括：

获取待识别图像，其中，所述待识别图像中包括文本；

基于预先训练的文本识别模型对所述待识别图像进行文本识别，得到所述待识别图像中的文本内容；

其中，所述文本识别模型是基于如权利要求1-9任一项所述的方法得到的。

11. 根据权利要求10所述的方法，其中，基于预先训练的文本识别模型对所述待识别图像进行文本识别，得到所述待识别图像中的文本内容，包括：

根据所述文本识别模型确定所述待识别图像的多模态特征图，并根据所述多模态特征图确定所述待识别图像中的文本内容；

其中，所述待识别图像的多模态特征图用于表征：所述待识别图像的视觉特征和语义特征。

12. 一种文本识别模型的训练装置，其特征在于，包括：

预测单元，用于对获取到的第一样本图像中的部分图像进行掩码预测，得到与所述第一样本图像对应的预测完整图像；

所述预测单元还用于，对获取到的第二样本图像中的部分文本进行所述掩码预测，得到与所述部分文本对应的预测文本内容，其中，所述第一样本图像和所述第二样本图像为不同的图像；

训练单元，用于根据所述预测完整图像和所述预测文本内容训练得到预训练模型；

生成单元，用于根据所述预训练模型生成文本识别模型，其中，所述文本识别模型用于对待识别图像进行文本识别。

13. 根据权利要求12所述的装置，其中，所述第一样本图像中的图像和所述第二样本图像中的文本所表征的内容不同。

14. 根据权利要求12所述的装置，其中，所述生成单元，包括：

获取子单元，用于获取待识别任务和训练图像，其中，所述训练图像中包括文本；

训练子单元，用于根据所述待识别任务和所述训练图像，对所述预训练模型进行训练，得到所述文本识别模型。

15. 根据权利要求14所述的装置,其中,所述训练子单元,包括:

输入模块,用于将所述训练图像输入至所述预训练模型,得到所述训练图像对应的多模态特征图,其中,所述多模态特征图用于表征所述训练图像对应的图像特征和语义特征;
生成模块,用于根据所述待识别任务和所述多模态特征图,生成所述文本识别模型。

16. 根据权利要求15所述的装置,其中,所述生成模块,包括:

第三预测子模块,用于根据所述多模态特征图,预测所述训练图像在所述待识别任务下的预测识别结果;

构建子模块,用于根据所述训练图像预设的真实识别结果、以及所述预测识别结果,构建所述文本识别模型。

17. 根据权利要求12-16中任一项所述的装置,其中,所述预测单元,包括:

遮盖子单元,用于随机遮盖目标对象中的部分对象;

预测子单元,用于根据所述目标对象中未被遮盖的对象,对所述目标对象中被遮盖的部分对象进行预测,得到预测结果;

其中,若所述目标对象为第一样本图像,则所述目标对象中的部分对象为部分图像,所述预测结果为所述预测完整图像;若所述目标对象为第二样本图像,则所述目标对象中的部分对象为部分文本,所述预测结果为所述预测文本内容。

18. 根据权利要求17所述的装置,其中,所述预测子单元,包括:

提取模块,用于提取所述目标对象中未被遮盖的对象对应的对象特征,得到第一对象特征;

预测模块,用于根据所述第一对象特征,对所述目标对象中被遮盖的部分对象进行预测,得到所述预测结果;

其中,若所述目标对象为第一样本图像,则所述第一对象特征为第一视觉特征;若所述目标对象为第二样本图像,则所述第一对象特征为第一语义特征。

19. 根据权利要求18所述的装置,其中,所述目标对象为第一样本图像,所述第一对象特征为第一视觉特征;所述预测模块,包括:

第一预测子模块,用于根据所述第一视觉特征,预测所述第一样本图像中被遮盖的部分图像对应的视觉特征第二视觉特征;

第一确定子模块,用于根据所述第二视觉特征,确定所述第一样本图像中被遮盖的部分图像;

第一生成子模块,用于根据所述第一样本图像中未被遮盖的图像、以及确定出的所述第一样本图像中被遮盖的部分图像,生成所述预测完整图像。

20. 根据权利要求18或19所述的装置,其中,所述目标对象为第二样本图像,所述第一对象特征为第一语义特征;所述预测模块,包括:

第二预测子模块,用于根据所述第一语义特征,预测所述第二样本图像中被遮盖的部分文本对应的语义特征,得到第二语义特征;

第二生成子模块,用于根据所述第二语义特征,生成所述预测文本内容。

21. 一种文本识别装置,其特征在于,包括:

获取单元,用于获取待识别图像,其中,所述待识别图像中包括文本;

识别单元,用于基于预先训练的文本识别模型对所述待识别图像进行文本识别,得到

所述待识别图像中的文本内容；

其中，所述文本识别模型是基于如权利要求1-9任一项所述的方法得到的。

22. 根据权利要求21所述的装置，其中，所述识别单元，包括：

第一确定单元，用于根据所述文本识别模型确定所述待识别图像的多模态特征图；

第二确定单元，用于根据所述多模态特征图确定所述待识别图像中的文本内容；

其中，所述待识别图像的多模态特征图用于表征：所述待识别图像的视觉特征和语义特征。

23. 一种电子设备，其特征在于，包括：

至少一个处理器；以及

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行权利要求1-9任一项所述的方法；或者，以使所述至少一个处理器能够执行权利要求10或11所述的方法。

24. 一种存储有计算机指令的非瞬时计算机可读存储介质，其特征在于，所述计算机指令用于使所述计算机执行根据权利要求1-9任一项所述的方法；或者，所述计算机指令用于使所述计算机执行根据权利要求10或11所述的方法。

文本识别模型的训练方法、文本识别方法及装置

技术领域

[0001] 本公开涉及人工智能 (Artificial Intelligence, AI) 技术领域, 具体为深度学习、计算机视觉技术领域, 可应用于光学字符识别 (Optical Character Recognition, OCR) 等场景, 尤其涉及一种文本识别模型的训练方法、文本识别方法及装置。

背景技术

[0002] OCR技术在教育、金融、医疗、交通和保险等各行各业都取得了广泛的关注和应用。

[0003] 在相关技术中, 可以结合OCR技术和深度学习构建文本识别模型, 以基于文本识别模型对图像进行文本识别。

[0004] 然而, 文本识别模型通常依赖于视觉信息, 以基于视觉信息辨别图像中的文本内容, 存在识别的准确性偏低的弊端。

发明内容

[0005] 本公开提供了一种用于提高文本识别的可靠性的文本识别模型的训练方法、文本识别方法及装置。

[0006] 根据本公开的第一方面, 提供了一种文本识别模型的训练方法, 包括:

[0007] 对获取到的第一样本图像中的部分图像进行掩码预测, 得到与所述第一样本图像对应的预测完整图像;

[0008] 对获取到的第二样本图像中的部分文本进行所述掩码预测, 得到与所述部分文本对应的预测文本内容;

[0009] 根据所述预测完整图像和所述预测文本内容训练得到预训练模型, 并根据所述预训练模型生成文本识别模型, 其中, 所述文本识别模型用于对待识别图像进行文本识别。

[0010] 根据本公开的第二方面, 提供了一种文本识别方法, 包括:

[0011] 获取待识别图像, 其中, 所述待识别图像中包括文本;

[0012] 基于预先训练的文本识别模型对所述待识别图像进行文本识别, 得到所述待识别图像中的文本内容;

[0013] 其中, 所述文本识别模型是基于如第一方面所述的方法得到的。

[0014] 根据本公开的第三方面, 提供了一种文本识别模型的训练装置, 包括:

[0015] 预测单元, 用于对获取到的第一样本图像中的部分图像进行掩码预测, 得到与所述第一样本图像对应的预测完整图像;

[0016] 所述预测单元还用于, 对获取到的第二样本图像中的部分文本进行所述掩码预测, 得到与所述部分文本对应的预测文本内容;

[0017] 训练单元, 用于根据所述预测完整图像和所述预测文本内容训练得到预训练模型;

[0018] 生成单元, 用于根据所述预训练模型生成文本识别模型, 其中, 所述文本识别模型用于对待识别图像进行文本识别。

- [0019] 根据本公开的第四方面,提供了一种文本识别装置,包括:
- [0020] 获取单元,用于获取待识别图像,其中,所述待识别图像中包括文本;
- [0021] 识别单元,用于基于预先训练的文本识别模型对所述待识别图像进行文本识别,得到所述待识别图像中的文本内容;
- [0022] 其中,所述文本识别模型是基于如第一方面所述的方法得到的。
- [0023] 根据本公开的第五方面,提供了一种电子设备,包括:
- [0024] 至少一个处理器;以及
- [0025] 与所述至少一个处理器通信连接的存储器;其中,
- [0026] 所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行第一方面或第二方面所述的方法。
- [0027] 根据本公开的第六方面,提供了一种存储有计算机指令的非瞬时计算机可读存储介质,其中,所述计算机指令用于使所述计算机执行根据第一方面或第二方面所述的方法。
- [0028] 根据本公开的第七方面,提供了一种计算机程序产品,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可读存储介质中,电子设备的至少一个处理器可以从所述可读存储介质读取所述计算机程序,所述至少一个处理器执行所述计算机程序使得电子设备执行第一方面或第二方面所述的方法。
- [0029] 根据本公开的基于掩码预测得到第一样本图像对应的预测完整图像,基于掩码预测得到第二样本图像中的部分文本的预测文本内容,并结合预测完整图像和预测文本内容生成预训练模型,以基于预训练模型生成文本识别模型的技术方案,使得预训练模型学习到较强的图像视觉推理能力和文本语义推理能力,从而当基于预训练模型生成的文本识别模型进行文本识别时,提高文本识别的准确性和可靠性。
- [0030] 应当理解,本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征,也不用于限制本公开的范围。本公开的其他特征将通过以下的说明书而变得容易理解。

附图说明

- [0031] 附图用于更好地理解本方案,不构成对本公开的限定。其中:
- [0032] 图1是根据本公开第一实施例的示意图;
- [0033] 图2是根据本公开第二实施例的示意图;
- [0034] 图3是根据本公开第三实施例的示意图;
- [0035] 图4是根据本公开第四实施例的示意图;
- [0036] 图5是根据本公开第五实施例的示意图;
- [0037] 图6是根据本公开第六实施例的示意图;
- [0038] 图7是根据本公开第七实施例的示意图;
- [0039] 图8是根据本公开第八实施例的示意图;
- [0040] 图9是根据本公开第九实施例的示意图;
- [0041] 图10是用来实现本公开实施例的文本识别模型的训练方法、文本识别方法的电子设备的框图。

具体实施方式

[0042] 以下结合附图对本公开的示范性实施例做出说明,其中包括本公开实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本公开的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0043] 在结合OCR技术和深度学习构建文本识别模型时,可以采用“模块分离”的方式实现,也可以采用“端到端模型”的方式。

[0044] 示例性的,“模块分离”的方式是指,构建文本检测模块、信息抽取模块、文本识别模块,以结合三个模块构建文本识别模型。

[0045] 若采用“模块分离”的方式,需要预先构建各模块,并将各模块进行结合,过程相对繁琐,效率相对偏低,且准确性会累积叠加,造成基于该方式构建的文本识别模型的识别准确性偏低的弊端。

[0046] 示例性的,“端到端模型”的方式是指,从输入端到输出端会得到一个预测结果,如在输入端输入图像,在输出端对得到图像的预测文本内容。

[0047] 然而,采用“端到端模型”的方式需要进行数据标注,如对图像的真实文本内容进行标注,且用于提供训练的数据比较有效,导致训练得到的文本识别模型的可靠性偏低的弊端。

[0048] 且基于上述任一方法训练得到的文本识别模型,通常只做二类判断,不同的垂类有不同类别字段需求的时候,需要重新设计文本识别模型,尤其是分类的通道数,文本识别模型也要重新训练,无法复用。

[0049] 例如,OCR技术中的图像文字检测模型(EAST)、分割的文字检测模型(DB)、和文本检测器(LOMO)等,一般只能用作两类判断,如文档(text)类和非文档类(non-text)。若需要解决某个具体垂类下用户感兴趣的字段的识别需求,需要增加分类类别数。

[0050] 在一些实施例中,可以通过检测扩增类别的方式,训练得到新的文本识别模型,可以通过在原有文本识别模型的基础上,附加额外的语言模型做字段分类。

[0051] 例如,若文本识别模型为OCR技术中的端到端文本检测与识别(FOTS)和文本检测加识别模型(Mask Text Spotter),则需要附加额外的语言模型如双向编码器表示(Bidirectional Encoder Representations,BERT),从而得到新的文本识别模型,且由于增加了额外的语言模型,则需要增加额外的训练,从而导致训练成本偏高,效率偏低等弊端。

[0052] 为了避免上述技术问题中的至少一种,本公开的发明人经过创造性地劳动,得到了本公开的发明构思:采用“端到端模型”的方式训练得到预训练模型,即对模型底座进行端到端的预训练,且结合视觉维度和语义维度进行预训练,以基于预训练得到的底座生成文本识别模型。

[0053] 基于上述发明构思,本公开提供一种文本识别模型的训练方法、文本识别方法及装置,涉及人工智能技术领域,具体为深度学习、计算机视觉技术领域,可应用于OCR等场景,以达到提高文本识别模型对文本识别的可靠性。

[0054] 图1是根据本公开第一实施例的示意图,如图1所示,本实施例提供的文本识别模型的训练方法,包括:

[0055] S101:对获取到的第一样本图像中的部分图像进行掩码预测,得到与第一样本图像对应的预测完整图像。

[0056] 示例性的,本实施例的执行主体可以为文本识别模型的训练装置(下文简称为训练装置),训练装置可以为服务器(如云端服务器,或者,本地服务器,或者,服务器集群),也可以为终端设备,也可以为计算机,也可以为处理器,也可以为芯片等,本实施例不做限定。

[0057] 其中,掩码预测是指,对部分图像或文本等进行掩码(mask)处理(或者,称为遮盖处理),并还原mask处理前,即遮盖处理前的图像或文本等的完整的图像或文本等。

[0058] 相应的,该步骤可以理解为:获取包括文本的第一样本图像,对第一样本图像的部分图像进行mask处理,并基于mask处理后的图像预测完整的第一样本图像(即预测完整图像)。

[0059] 也就是说,该步骤可以理解为图像重建任务(mask image modelling),以结合掩码预测的方式对第一样本图像进行图像重建。

[0060] S102:对获取到的第二样本图像中的部分文本进行掩码预测,得到与部分文本对应的预测文本内容。

[0061] 结合上述分析,该步骤可以理解为:获取包括文本的第二样本图像,对第二样本图像中的部分文本进行mask处理,并基于mask处理后的文本预测被mask处理的部分文本的文本内容(即预测文本内容)。

[0062] 也就是说,该步骤可以理解为文本重建任务(mask OCR modelling),以结合掩码预测的方式对第二样本图像进行文本重建,且具体为对第二样本图像中的部分文本进行重建。

[0063] 值得说明的是,第一样本图像和第二样本图像可以为相同的图像,也可以为不同的图像,本实施例不做限定。

[0064] S103:根据预测完整图像和预测文本内容训练得到预训练模型,并根据预训练模型生成文本识别模型。

[0065] 其中,文本识别模型用于对待识别图像进行文本识别。

[0066] 预训练模型可以理解为文本识别模型的底座,或者,可以理解为文本识别模型的隐藏层。

[0067] 结合上述分析可知,预训练模型是基于图像重建和文本重建而训练得到的,以使得预训练模型学习到较强的图像视觉推理能力和文本语义推理能力,以使得基于预训练模型生成的文本识别模型具有较强的准确性和可靠性。

[0068] 且在本实施例中,可以实现端到端的模型训练,即可以直接基于第一样本图像和第二样本图像输出各自对应的预测结果,如第一样本图像对应的预测结果为预测完整图像,第二样本图像对应的预测结果为预测文本内容,无需增加其他环节,如基于人工或OCR技术对第二样本图像进行文本检测,以得到文本的环节,从而提高了训练效率,节约训练资源和成本。

[0069] 基于上述分析可知,本公开实施例提供了一种文本识别模型的训练方法,包括:对获取到的第一样本图像中的部分图像进行掩码预测,得到与第一样本图像对应的预测完整图像,对获取到的第二样本图像中的部分文本进行掩码预测,得到与部分文本对应的预测文本内容,根据预测完整图像和预测文本内容训练得到预训练模型,并根据预训练模型生

成文本识别模型,其中,文本识别模型用于对待识别图像进行文本识别,在本实施例中,通过基于掩码预测得到第一样本图像对应的预测完整图像,基于掩码预测得到第二样本图像中的部分文本的预测文本内容,并结合预测完整图像和预测文本内容生成预训练模型,以基于预训练模型生成文本识别模型的技术特征,使得预训练模型学习到较强的图像视觉推理能力和文本语义推理能力,从而当基于预训练模型生成的文本识别模型进行文本识别时,提高文本识别的准确性和可靠性。

[0070] 图2是根据本公开第二实施例的示意图,如图2所示,本实施例提供的文本识别模型的训练方法,包括:

[0071] S201:获取目标对象。

[0072] 其中,目标对象包括第一样本图像和第二样本图像。

[0073] 应该理解的是,为了避免繁琐的陈述,关于本实施例与上述实施例相同的技术特征,本实施例不做赘述。

[0074] S202:随机遮盖目标对象中的部分对象,并根据目标对象中未被遮盖的对象,对目标对象中被遮盖的部分对象进行预测,得到预测结果。

[0075] 其中,若目标对象为第一样本图像,则目标对象中的部分对象为部分图像,预测结果为预测完整图像。

[0076] 若目标对象为第二样本图像,则目标对象中的部分对象为部分文本,预测结果为预测文本内容。

[0077] 在一些实施例中,根据目标对象中未被遮盖的对象,对目标对象中被遮盖的部分对象进行预测,得到预测结果包括如下步骤:

[0078] 第一步骤:提取目标对象中未被遮盖的对象对应的对象特征,得到第一对象特征。

[0079] 第二步骤:根据第一对象特征,对目标对象中被遮盖的部分对象进行预测,得到预测结果。

[0080] 其中,若目标对象为第一样本图像,则第一对象特征为第一视觉特征。若目标对象为第二样本图像,则第一对象特征为第一语义特征。

[0081] S203:根据预测完整图像和预测文本内容训练得到预训练模型,并根据预训练模型生成文本识别模型。

[0082] 其中,文本识别模型用于对待识别图像进行文本识别。

[0083] 为使读者更加深刻地理解本公开的实现原理,现结合图3对上述实施例(如图1和图2所示的实施例)进行详细地阐述。

[0084] 图3是根据本公开第三实施例的示意图,如图3所示,本实施例提供的文本识别模型的训练方法,包括:

[0085] S301:获取第一样本图像。

[0086] 同理,为了避免繁琐的陈述,关于本实施例与上述实施例相同的技术特征,本实施例不做赘述。

[0087] S302:随机遮盖第一样本图像中的部分图像。

[0088] 应该理解的是,网络模型的训练通常是迭代训练的过程,在本实施例中,由于每一次迭代训练都是随机遮盖第一样本图像的部分图像,因此,第一样本图像的数量可以为一个,当然,第一样本图像的数量也可以为多个,本实施例不做限定。

[0089] S303:根据第一样本图像中未被遮盖的图像,对第一样本图像中被遮盖的部分图像进行预测,得到预测完整图像。

[0090] 示例性地,在对第一样本图像进行随机遮盖之后,第一样本图像中的部分图像是被遮盖的,另外部分图像是未被遮盖的,则可以基于未被遮盖的图像确定完整的第一样本图像(即预测完整图像)。

[0091] 在本实施例中,通过结合“随机遮盖+预测”的方式,确定预测完整图像,可以增加训练过程中的不可确定性,从而提高训练得到的预训练模型还原完整图像的可靠性。

[0092] 其中,S302-S303可以基于掩膜自编码器(MAE)实现。也就是说,可以将第一样本图像输入至掩膜自编码器,输出预测完整图像。

[0093] 在一些实施例中,S303可以包括如下步骤:

[0094] 第一步骤:提取第一样本图像中未被遮盖的图像对应的视觉特征,得到第一视觉特征。

[0095] 其中,视觉特征包括如纹理特征、轮廓特征、颜色特征、以及形状特征等,此处不再一一列举。

[0096] 相应的,第一视觉特征,是指第一样本图像中未被遮盖的图像对应的如纹理特征、轮廓特征、颜色特征、以及形状特征等。

[0097] 第二步骤:根据第一视觉特征,对第一样本图像中被遮盖的部分图像进行预测,得到预测完整图像。

[0098] 在本实施例中,通过结合未被遮盖的图像对应的如纹理特征、轮廓特征、颜色特征、以及形状特征等视觉特征,得到预测完整图像,相当于基于视觉上下文得到预测完整图像,以训练得到能够完成视觉线索的上下文知识学习的预训练模型。

[0099] 在一些实施例中,第二步骤可以包括如下子步骤:

[0100] 第一子步骤:根据第一视觉特征,预测第一样本图像中被遮盖的部分图像对应的视觉特征,得到第二视觉特征。

[0101] 示例性的,结合上述分析,该子步骤可以理解为:根据未被遮盖的图像对应的如纹理特征、轮廓特征、颜色特征、以及形状特征等视觉特征,预测得到被遮盖的部分图像对应的如纹理特征、轮廓特征、颜色特征、以及形状特征等视觉特征。

[0102] 第二子步骤:根据第二视觉特征,确定第一样本图像中被遮盖的部分图像。

[0103] 示例性的,在得到被遮盖的部分图像对应的如纹理特征、轮廓特征、颜色特征、以及形状特征等视觉特征后,可以基于该视觉特征补充和修复被遮盖的部分图像。

[0104] 第三子步骤:根据第一样本图像中未被遮盖的图像、以及确定出的第一样本图像中被遮盖的部分图像,生成预测完整图像。

[0105] 结合上述分析,在对被遮盖的部分图像进行补充和修复后,即还原了被遮盖的部分图像,将未被遮盖的部分图像和还原的被遮盖的部分图像进行拼接,从而得到预测完整图像,即还原第一样本图像,以使得预测完整图像与第一样本图像高度贴合,提高预测完整图像的准确性和可靠性。

[0106] S304:获取第二样本图像。

[0107] 结合上述分析可知,第一样本图像和第二样本图像可以为相同的图像,相应的,若第一样本图像和第二样本图像为相同的图像,则可以省略该步骤。

[0108] S305:随机遮盖第二样本图像中的部分文本。

[0109] 同理,网络模型的训练通常是迭代训练的过程,在本实施例中,由于每一次迭代训练都是随机遮盖第二样本图像的部分文本,因此,第二样本图像的数量可以为一个,当然,第二样本图像的数量也可以为多个,本实施例不做限定。

[0110] 例如,可以随机遮盖第二样本图像中的部分词,或者部分句子等。

[0111] S306:根据第二样本图像中未被遮盖的文本,对第二样本图像中被遮盖的部分文本进行预测,得到预测文本内容。

[0112] 示例性地,在对第二样本图像进行随机遮盖之后,第二样本图像中的部分文本是被遮盖的,另外部分文本是未被遮盖的,则可以基于未被遮盖的文本确定被遮盖的部分文本的文本内容(即预测文本内容)。

[0113] 在本实施例中,通过结合“随机遮盖+预测”的方式,确定文本内容,可以增加训练过程中的不可确定性,从而提高训练得到的预训练模型还原完整图像的可靠性。

[0114] 其中,S305-S306可以基于掩码语言模型(Masked Language Model,MLM)实现。也就是说,可以将第二样本图像输入至掩码语言模型,输出预测文本内容。

[0115] 在一些实施例中,S306可以包括如下步骤:

[0116] 第一步骤:提取第二样本图像中未被遮盖的文本对应的语义特征,得到第一语义特征。

[0117] 其中,语义特征是指各字符串之间的逻辑关系的特征。相应的,第一语义特征可以理解为,未被遮盖的文本中包括的各字符串之间的逻辑关系的特征,也可以理解为未被遮盖的文本中的各文字(字和/或词)之间的关联关系的特征。

[0118] 第二步骤:根据第一语义特征,对第二样本图像中被遮盖的部分文本进行预测,得到预测文本内容。

[0119] 在本实施例中,通过结合未被遮盖的文本对应的各字符串之间的逻辑关系等视觉特征,得到预测文本内容,相当于基于语义上下文得到预测文本内容,以训练得到能够完成语义线索的上下文知识学习的预训练模型。

[0120] 在一些实施例中,第二步骤可以包括如下子步骤:

[0121] 第一子步骤:根据第一语义特征,预测第二样本图像中被遮盖的部分文本对应的语义特征,得到第二语义特征。

[0122] 示例性的,结合上述分析,该子步骤可以理解为:根据未被遮盖的文本对应的如各字符串之间的逻辑关系的特征等语义特征,预测得到被遮盖的部分文本对应的如各字符串之间的逻辑关系的特征等语义特征。

[0123] 第二子步骤:根据第二语义特征,生成预测文本内容。

[0124] 示例性的,在得到未被遮盖的文本对应的如各字符串之间的逻辑关系的特征等语义特征后,可以基于该语义特征补充和修复被遮盖的部分文本的语义特征。

[0125] 结合上述分析,在对被遮盖的部分文本的语义特征进行补充和修复后,即还原了被遮盖的部分文本的语义特征,可以确定该语义特征对应的文本内容(即预测文本内容),以使得预测文本内容与被遮盖的部分文本的文本内容高度贴合,提高预测文本内容的准确性和可靠性。

[0126] S307:根据预测完整图像和预测文本内容训练得到预训练模型,并根据预训练模

型生成文本识别模型。

[0127] 其中,文本识别模型用于对待识别图像进行文本识别。

[0128] 图4是根据本公开第四实施例的示意图,如图4所示,本实施例提供的文本识别模型的训练方法,包括:

[0129] S401:对获取到的第一样本图像中的部分图像进行掩码预测,得到与第一样本图像对应的预测完整图像。

[0130] 同理,为了避免繁琐的陈述,关于本实施例与上述实施例相同的技术特征,本实施例不做赘述。

[0131] S402:对获取到的第二样本图像中的部分文本进行掩码预测,得到与部分文本对应的预测文本内容。

[0132] S403:根据预测完整图像和预测文本内容训练得到预训练模型。

[0133] 示例性的,可以基于预测完整图像和预测文本内容,对基础网络模型进行训练,以得到预训练模型。

[0134] 例如,可以基于预测完整图像和预测文本内容,对基础网络模型的模型参数进行调整,以得到预训练模型。

[0135] 其中,基础网络模型可以为视觉转换器(Vision Transformer,ViT),也可以为神经网络模型(Backbone),如卷积神经网络模型(CNN),也可以为其他网络模型,本实施例不做限定。

[0136] S404:获取待识别任务和训练图像。

[0137] 其中,训练图像中包括文本。

[0138] 其中,待识别任务可以为基于文本识别模型的识别需求确定的,如待识别任务可以为文字检测任务,也可以为文本识别任务,也可以为字段分类任务,也可以为其他识别任务,此处不再一一列举。

[0139] S405:根据待识别任务和训练图像,对预训练模型进行训练,得到文本识别模型。

[0140] 其中,文本识别模型用于对待识别图像进行文本识别。

[0141] 结合上述分析可知,预训练模型既有完成视觉线索的上下文知识学习的模型,又具有语义线索上下文知识学习的模型,即预训练模型为多模态特征提取底座,因此,当结合预训练模型训练得到的文本识别模型,既具有基于视觉线索的上下文知识识别能力,又具有基于语义线索上下文知识识别能力。

[0142] 且通过结合待识别任务对预训练模型进行训练,可以实现基于不同的识别需求训练得到与之对应的文本识别模型,从而提高训练得到文本识别模型的灵活性和多样性,可以广泛地应用于各种识别场景,满足不同的识别需求。

[0143] 在一些实施例中,可以将预训练模型(即多模态特征提取底座)加载至文本检测网络模型(Efficient and Accuracy Scene Text,EAST)、基于分割的文字检测网络(Differentiable Binarization,DB)、文本检测网络(Look More Than Once,LOMO)等,以实现文本识别模型的文字检测任务;又如,可以将预训练模型加载至卷积循环神经网络(Convolutional Recurrent Neural Network,CRNN),其中,卷积循环神经网络可以采用联结主义时间分类(Connectionist Temporal Classification,CTC)解码方式,也可以采用注意机制(Attention)解码方式,也可以采用转换器(transformer)解码方法等,以实现文

本识别模型的文本识别任务；再如，可以将预训练模型加载至全连接网络模型 (Fully Connected, FC)，或者卷积神经网络模型 (Convolutional Neural Networks, CNN)，以实现文本识别模型的字段分类任务。

[0144] 在一些实施例中，S405可以包括如下步骤：

[0145] 第一步骤：将训练图像输入至预训练模型，得到训练图像对应的多模态特征图 (Multi-modal Feature Maps)。

[0146] 结合上述分析，多模态特征图用于表征训练图像的多个维度的特征，如视觉维度的特征和语义维度的特征。如多模态特征图可以用于表征训练图像对应的图像特征和语义特征。

[0147] 在一些实施例中，多模态特征图可以表示成 $(d \times h \times w)$ ，其中， d 表示特征通道数， h 和 w 表示多模态特征图的高和宽。

[0148] 第二步骤：根据待识别任务和多模态特征图，生成文本识别模型。

[0149] 在本实施例中，由于多模态特征图可以从多个维度对训练图像的特征进行表征，既可以表征训练图像的视觉特征，又可以表征训练图像的语义特征，且表征的视觉特征和语义特征具有较强的可靠性和全面性，因此，结合多模态特征图生成的文本识别模型具有较强的可靠性和准确性。

[0150] 在一些实施例中，第二步骤可以包括如下子步骤：

[0151] 第一子步骤：根据多模态特征图，预测训练图像在待识别任务下的预测识别结果。

[0152] 示例性的，可以将多模态特征图输入至卷积循环神经网络，得到预测识别结果 (如预测文本结果)。

[0153] 第二子步骤：根据训练图像预设的真实识别结果、以及预测识别结果，构建文本识别模型。

[0154] 其中，真实识别结果可以为预先对训练图像进行标注得到的，标注的方式本实施例不做限定，如可以为人工标注的方式，也可以为自动标注的方式。

[0155] 示例性的，可以计算真实识别结果与预测识别结果之间的损失值，若损失值大于 (或等于) 预设的损失阈值，则迭代进行训练，反之，若损失值小于预设的损失阈值，则文本识别模型构建完成，或者，若迭代次数达到预设迭代次数，则文本识别模型构建完成。

[0156] 例如，若需要训练用于对火车票进行文本识别的文本识别模型，则训练图像为火车票图像，将火车票图像输入至预训练模型，输出火车票图像的多模态特征图，将多模态特征图输入至如卷积循环神经网络，输出如火车票图像中的“日期、车次、座位号”等预测识别结果，将该预测识别结果与预先标注的“日期、车次、座位号” (即真实识别结果) 进行比对，以训练得到文本识别模型，且训练得到的文本识别模型可以用于识别待识别的车票图像中的“日期、车次、座位号”文本内容。

[0157] 图5是根据本公开第五实施例的示意图，如图5所示，本实施例提供的文本识别方法，包括：

[0158] S501：获取待识别图像。

[0159] 其中，待识别图像中包括文本。

[0160] 示例性的，本实施例的执行主体可以为文本识别装置，文本识别装置可以为与训练装置相同的装置，也可以为与训练装置不同的装置，本实施例不做限定。

[0161] S502:基于预先训练的文本识别模型对待识别图像进行文本识别,得到待识别图像中的文本内容。

[0162] 其中,文本识别模型是基于如上述任一实施例所述的文本识别模型的训练方法得到的。

[0163] 在一些实施例中,S502可以包括如下步骤:

[0164] 第一步骤:根据文本识别模型确定待识别图像的多模态特征图。

[0165] 第二步骤:根据多模态特征图确定待识别图像中的文本内容。

[0166] 其中,待识别图像的多模态特征图用于表征:待识别图像的视觉特征和语义特征。

[0167] 示例性的,结合上述分析,文本识别模型包括预训练模型,若文本识别模型是通过将预训练模型加载至卷积循环神经网络训练得到的,即文本识别模型还包括卷积循环神经网络,则本实施例可以理解为:

[0168] 将待识别图像输入至预训练模型,输出多模态特征图,将多模态特征图输入至卷积循环神经网络,输出待识别图像中的文本内容。

[0169] 图6是根据本公开第六实施例的示意图,如图6所示,本实施例提供的文本识别模型的训练装置600,包括:

[0170] 预测单元601,用于对获取到的第一样本图像中的部分图像进行掩码预测,得到与第一样本图像对应的预测完整图像。

[0171] 预测单元601,还用于对获取到的第二样本图像中的部分文本进行掩码预测,得到与部分文本对应的预测文本内容。

[0172] 训练单元602,用于根据预测完整图像和预测文本内容训练得到预训练模型。

[0173] 生成单元603,用于根据预训练模型生成文本识别模型,其中,文本识别模型用于对待识别图像进行文本识别。

[0174] 图7是根据本公开第七实施例的示意图,如图7所示,本实施例提供的文本识别模型的训练装置700,包括:

[0175] 预测单元701,用于对获取到的第一样本图像中的部分图像进行掩码预测,得到与第一样本图像对应的预测完整图像。

[0176] 预测单元701还用于,对获取到的第二样本图像中的部分文本进行掩码预测,得到与部分文本对应的预测文本内容。

[0177] 结合图7,在一些实施例中,预测单元701,包括:

[0178] 遮盖子单元7011,用于随机遮盖目标对象中的部分对象。

[0179] 预测子单元7012,用于根据目标对象中未被遮盖的对象,对目标对象中被遮盖的部分对象进行预测,得到预测结果。

[0180] 其中,若目标对象为第一样本图像,则目标对象中的部分对象为部分图像,预测结果为预测完整图像;若目标对象为第二样本图像,则目标对象中的部分对象为部分文本,预测结果为预测文本内容。

[0181] 在一些实施例中,预测子单元7012,包括:

[0182] 提取模块,用于提取目标对象中未被遮盖的对象对应的对象特征,得到第一对象特征。

[0183] 预测模块,用于根据第一对象特征,对目标对象中被遮盖的部分对象进行预测,得

到预测结果。

[0184] 其中,若目标对象为第一样本图像,则第一对象特征为第一视觉特征;若目标对象为第二样本图像,则第一对象特征为第一语义特征。

[0185] 在一些实施例中,目标对象为第一样本图像,所述第一对象特征为第一视觉特征;预测模块,包括:

[0186] 第一预测子模块,用于根据第一视觉特征,预测第一样本图像中被遮盖的部分图像对应的视觉特征,得到第二视觉特征。

[0187] 第一确定子模块,用于根据第二视觉特征,确定第一样本图像中被遮盖的部分图像。

[0188] 第一生成子模块,用于根据第一样本图像中未被遮盖的图像、以及确定出的第一样本图像中被遮盖的部分图像,生成预测完整图像。

[0189] 在一些实施例中,目标对象为第二样本图像,所述第一对象特征为第一语义特征;预测模块,包括:

[0190] 第二预测子模块,用于根据第一语义特征,预测第二样本图像中被遮盖的部分文本对应的语义特征,得到第二语义特征。

[0191] 第二生成子模块,用于根据第二语义特征,生成预测文本内容。

[0192] 训练单元702,用于根据预测完整图像和预测文本内容训练得到预训练模型。

[0193] 生成单元703,用于根据预训练模型生成文本识别模型,其中,文本识别模型用于对待识别图像进行文本识别。

[0194] 结合图7,在一些实施例中,生成单元703,包括:

[0195] 获取子单元7031,用于获取待识别任务和训练图像,其中,训练图像中包括文本。

[0196] 训练子单元7032,用于根据待识别任务和训练图像,对预训练模型进行训练,得到文本识别模型。

[0197] 在一些实施例中,训练子单元7032,包括:

[0198] 输入模块,用于将训练图像输入至预训练模型,得到训练图像对应的多模态特征图。

[0199] 生成模块,用于根据待识别任务和多模态特征图,生成文本识别模型。

[0200] 在一些实施例中,生成模块,包括:

[0201] 第三预测子模块,用于根据多模态特征图,预测训练图像在待识别任务下的预测识别结果。

[0202] 构建子模块,用于根据训练图像预设的真实识别结果、以及预测识别结果,构建本识别模型。

[0203] 图8是根据本公开第八实施例的示意图,如图8所示,本实施例提供的文本识别装置800,包括:

[0204] 获取单元801,用于获取待识别图像,其中,待识别图像中包括文本。

[0205] 识别单元802,用于基于预先训练的文本识别模型对待识别图像进行文本识别,得到待识别图像中的文本内容。

[0206] 其中,文本识别模型是基于如上任一实施例所述的文本识别模型的训练方法得到的。

[0207] 结合图8可知,在一些实施例中,识别单元802,包括:

[0208] 第一确定单元8021,用于根据文本识别模型确定待识别图像的多模态特征图。

[0209] 第二确定单元8022,用于根据多模态特征图确定待识别图像中的文本内容。

[0210] 其中,待识别图像的多模态特征图用于表征:待识别图像的视觉特征和语义特征。

[0211] 图9是根据本公开第九实施例的示意图,如图9所示,本公开中的电子设备900可以包括:处理器901和存储器902。

[0212] 存储器902,用于存储程序;存储器902,可以包括易失性存储器(英文:volatile memory),例如随机存取存储器(英文:random-access memory,缩写:RAM),如静态随机存取存储器(英文:static random-access memory,缩写:SRAM),双倍数据率同步动态随机存取存储器(英文:Double Data Rate Synchronous Dynamic Random Access Memory,缩写:DDR SDRAM)等;存储器也可以包括非易失性存储器(英文:non-volatile memory),例如快闪存储器(英文:flash memory)。存储器902用于存储计算机程序(如实现上述方法的应用程序、功能模块等)、计算机指令等,上述的计算机程序、计算机指令等可以分区存储在一个或多个存储器902中。并且上述的计算机程序、计算机指令、数据等可以被处理器901调用。

[0213] 上述的计算机程序、计算机指令等可以分区存储在一个或多个存储器902中。并且上述的计算机程序、计算机指令等可以被处理器901调用。

[0214] 处理器901,用于执行存储器902存储的计算机程序,以实现上述实施例涉及的方法中的各个步骤。

[0215] 具体可以参见前面方法实施例中的相关描述。

[0216] 处理器901和存储器902可以是独立结构,也可以是集成在一起的集成结构。当处理器901和存储器902是独立结构时,存储器902、处理器901可以通过总线903耦合连接。

[0217] 本实施例的电子设备可以执行上述方法中的技术方案,其具体实现过程和技术原理相同,此处不再赘述。

[0218] 本公开的技术方案中,所涉及的用户个人信息的收集、存储、使用、加工、传输、提供和公开等处理,均符合相关法律法规的规定,且不违背公序良俗。

[0219] 根据本公开的实施例,本公开还提供了一种电子设备、一种可读存储介质和一种计算机程序产品。

[0220] 根据本公开的实施例,本公开还提供了一种计算机程序产品,计算机程序产品包括:计算机程序,计算机程序存储在可读存储介质中,电子设备的至少一个处理器可以从可读存储介质读取计算机程序,至少一个处理器执行计算机程序使得电子设备执行上述任一实施例提供的方案。

[0221] 图10示出了可以用来实施本公开的实施例的示例电子设备1000的示意性框图。电子设备旨在表示各种形式的数字计算机,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字助理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例,并且不意在限制本文中描述的和/或者要求的本公开的实现。

[0222] 如图10所示,设备1000包括计算单元1001,其可以根据存储在只读存储器(ROM) 1002中的计算机程序或者从存储单元1008加载到随机访问存储器(RAM) 1003中的计算机程

序,来执行各种适当的动作和处理。在RAM 1003中,还可存储设备1000操作所需的各种程序和数据。计算单元1001、ROM 1002以及RAM 1003通过总线1004彼此相连。输入/输出(I/O)接口1005也连接至总线1004。

[0223] 设备1000中的多个部件连接至I/O接口1005,包括:输入单元1006,例如键盘、鼠标等;输出单元1007,例如各种类型的显示器、扬声器等;存储单元1008,例如磁盘、光盘等;以及通信单元1009,例如网卡、调制解调器、无线通信收发机等。通信单元1009允许设备1000通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据。

[0224] 计算单元1001可以是各种具有处理和计算能力的通用和/或专用处理组件。计算单元1001的一些示例包括但不限于中央处理单元(CPU)、图形处理单元(GPU)、各种专用的人工智能(AI)计算芯片、各种运行机器学习模型算法的计算单元、数字信号处理器(DSP)、以及任何适当的处理器、控制器、微控制器等。计算单元1001执行上文所描述的各个方法和处理,例如文本识别模型的训练方法、文本识别方法。例如,在一些实施例中,文本识别模型的训练方法、文本识别方法可被实现为计算机软件程序,其被有形地包含于机器可读介质,例如存储单元1008。在一些实施例中,计算机程序的部分或者全部可以经由ROM 1002和/或通信单元1009而被载入和/或安装到设备1000上。当计算机程序加载到RAM1003并由计算单元1001执行时,可以执行上文描述的文本识别模型的训练方法、文本识别方法的一个或多个步骤。备选地,在其他实施例中,计算单元1001可以通过其他任何适当的方式(例如,借助于固件)而被配置为执行文本识别模型的训练方法、文本识别方法。

[0225] 本文中以上描述的系统和技术和各种实施方式可以在数字电子电路系统、集成电路系统、场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、芯片上系统的系统(SOC)、复杂可编程逻辑设备(CPLD)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0226] 用于实施本公开的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器或控制器,使得程序代码当由处理器或控制器执行时使流程图和/或框图中所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行,作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务器上执行。

[0227] 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。

[0228] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0229] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)和互联网。

[0230] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。服务器可以是云服务器,又称为云计算服务器或云主机,是云计算服务体系中的一项主机产品,以解决了传统物理主机与VPS服务("Virtual Private Server",或简称"VPS")中,存在的管理难度大,业务扩展性弱的缺陷。服务器也可以为分布式系统的服务器,或者是结合了区块链的服务器。

[0231] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本发公开中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本公开公开的技术方案所期望的结果,本文在此不进行限制。

[0232] 上述具体实施方式,并不构成对本公开保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等,均应包含在本公开保护范围之内。

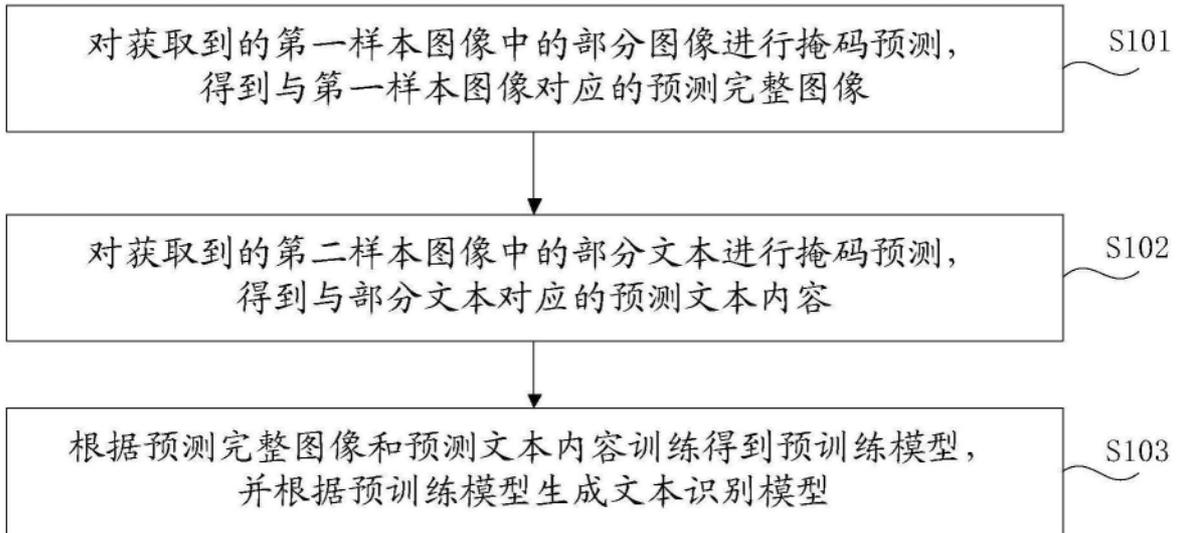


图1

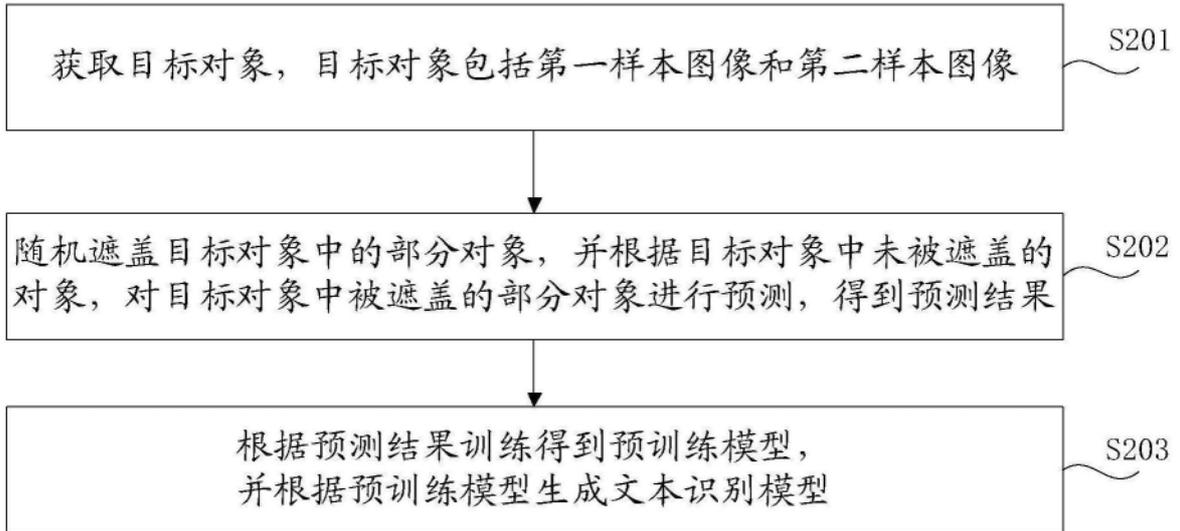


图2

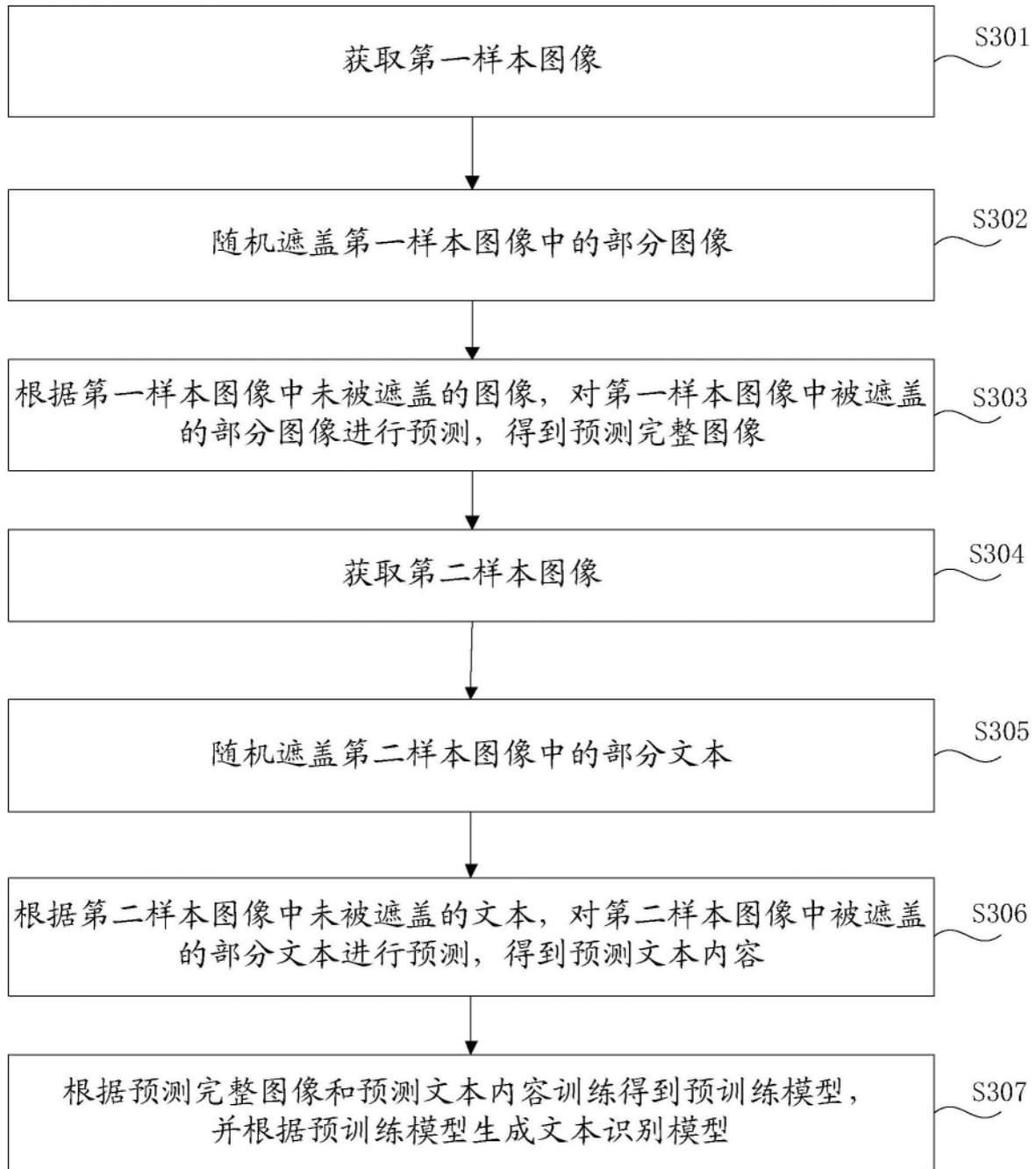


图3

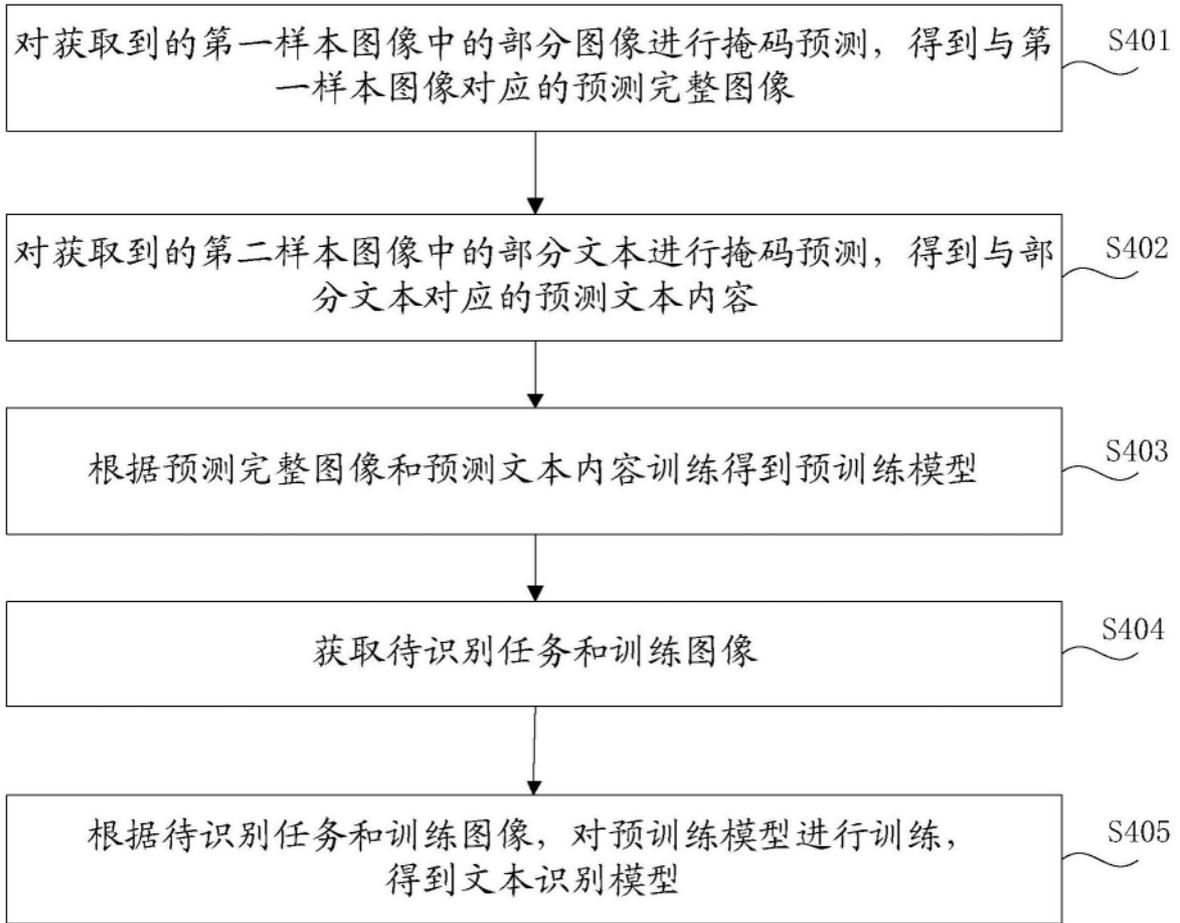


图4

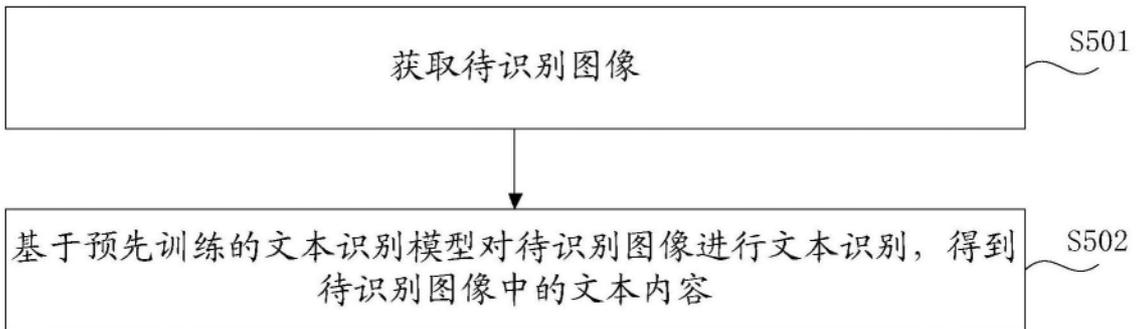


图5

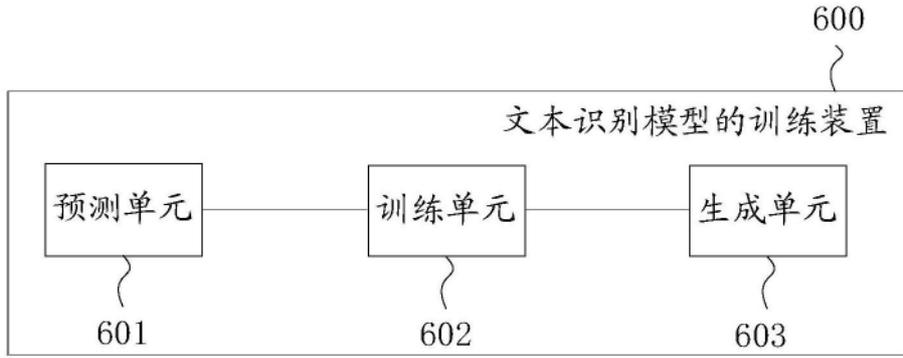


图6

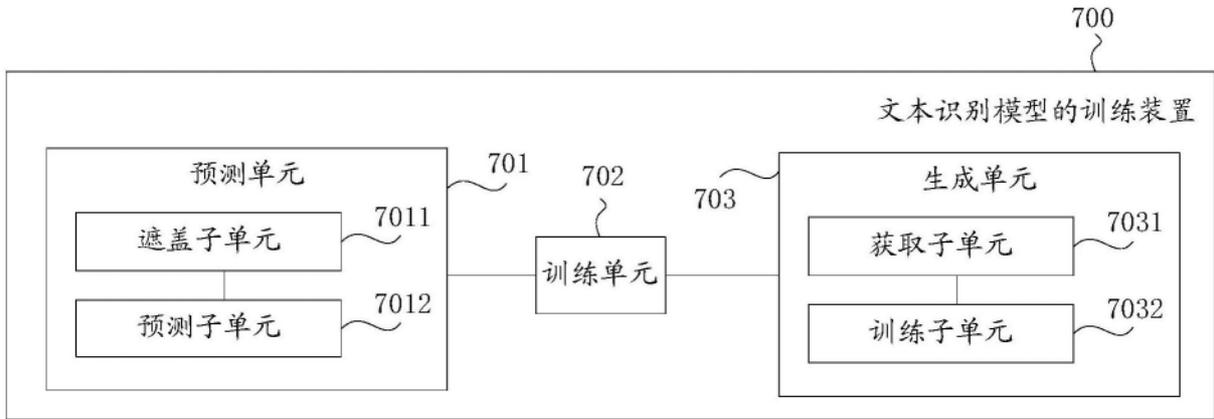


图7

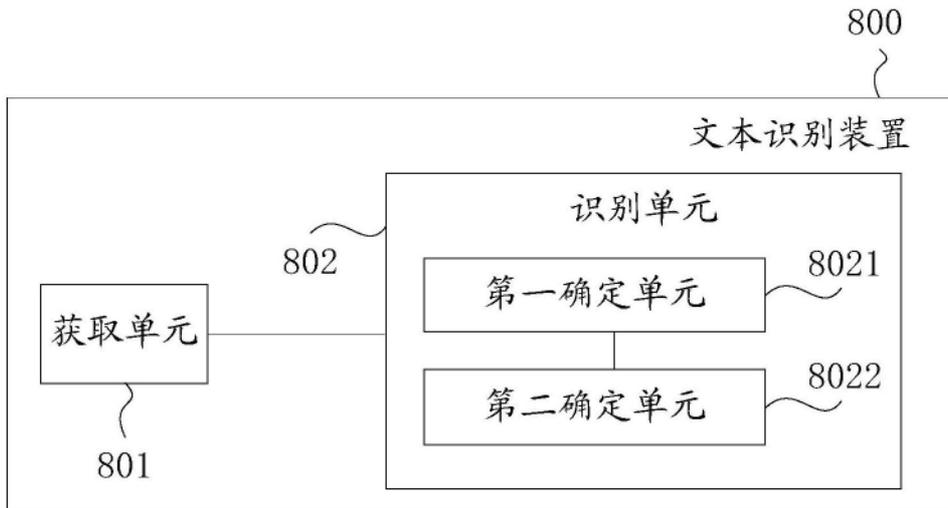


图8

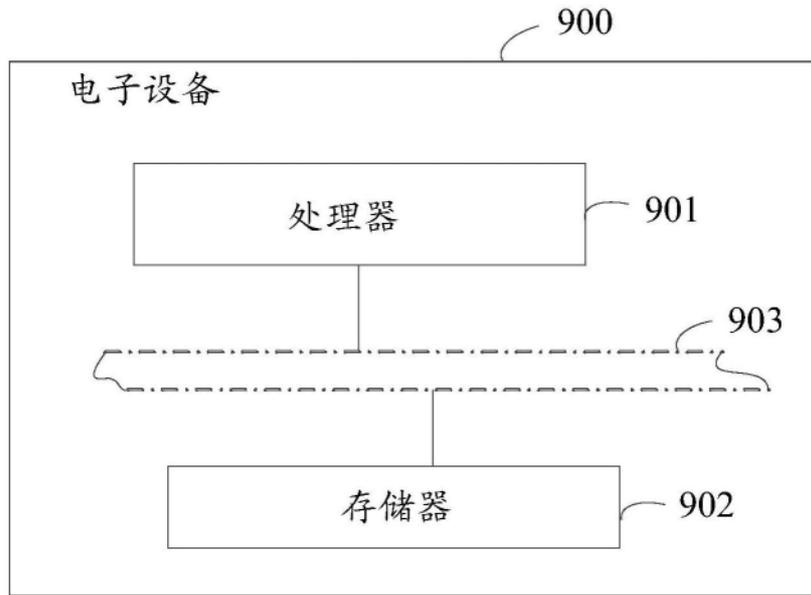


图9

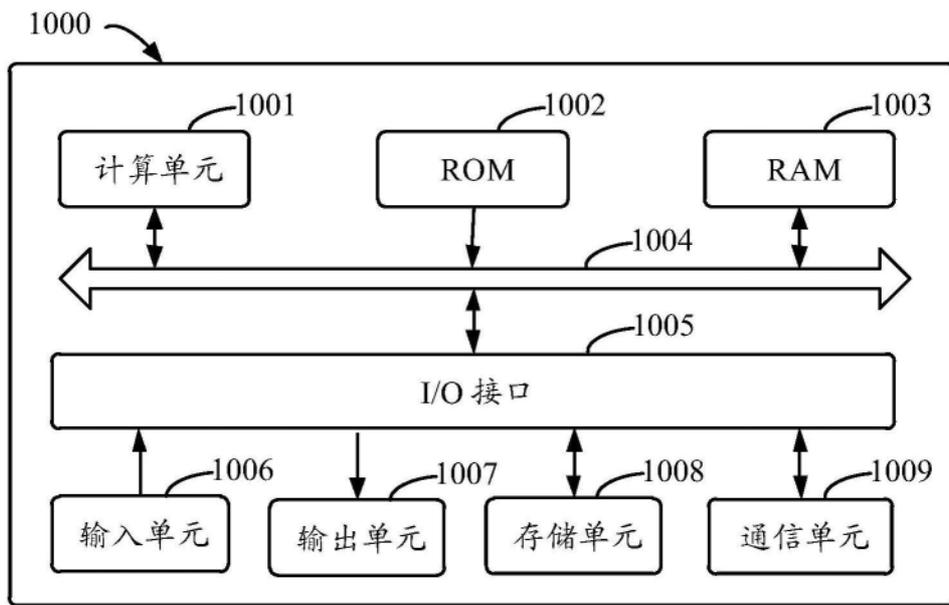


图10