



(12) 发明专利申请

(10) 申请公布号 CN 114398434 A

(43) 申请公布日 2022. 04. 26

(21) 申请号 202111528665.6 G06V 30/19 (2022.01)

(22) 申请日 2021.12.14 G06V 10/82 (2022.01)

(71) 申请人 北京百度网讯科技有限公司 G06K 9/62 (2022.01)

地址 100085 北京市海淀区上地十街10号 G06N 3/04 (2006.01)

百度大厦二层 G06N 3/08 (2006.01)

(72) 发明人 周侠 王乐义 李乔伊 秦铎浩
刘明浩

(74) 专利代理机构 北京清亦华知识产权代理事
务所(普通合伙) 11201

代理人 杜月

(51) Int. Cl.
G06F 16/25 (2019.01)
G06F 16/22 (2019.01)
G06F 16/583 (2019.01)
G06V 30/40 (2022.01)

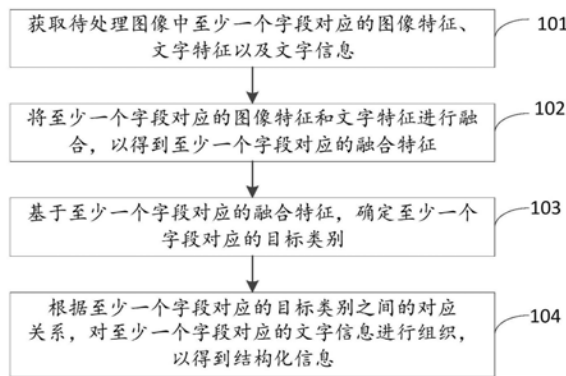
权利要求书5页 说明书23页 附图6页

(54) 发明名称

结构化信息抽取方法、装置、电子设备和存储介质

(57) 摘要

本公开提供了一种结构化信息抽取方法、装置、电子设备和存储介质,涉及人工智能技术领域,具体为深度学习、计算机视觉技术、光学字符识别OCR技术领域。具体实施方案为:获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息;将至少一个字段对应的图像特征和文字特征进行融合,以得到至少一个字段对应的融合特征;基于至少一个字段对应的融合特征,确定至少一个字段对应的目标类别;根据至少一个字段对应的目标类别之间的对应关系,对至少一个字段对应的文字信息进行组织,以得到结构化信息。由此,实现了将结构化信息抽取问题简化为分类问题,简化了结构化信息抽取的过程,并且,提高了结构化信息抽取的准确性。



1. 一种结构化信息抽取方法,包括:

获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息;

将至少一个所述字段对应的所述图像特征和所述文字特征进行融合,以得到至少一个所述字段对应的融合特征;

基于至少一个所述字段对应的融合特征,确定至少一个所述字段对应的目标类别;

根据至少一个所述字段对应的目标类别之间的对应关系,对至少一个所述字段对应的所述文字信息进行组织,以得到结构化信息。

2. 根据权利要求1所述的方法,至少一个所述字段对应的所述目标类别,为多个属性项类别中的至少一个或多个属性值类别中的至少一个;所述属性项类别和所述属性值类别之间具有一一对应关系;

其中,所述根据至少一个所述字段对应的目标类别之间的对应关系,对至少一个所述字段对应的文字信息进行组织,以得到结构化信息,包括:

根据所述目标类别为第一属性项类别的第一字段,查询所述目标类别为具有对应关系的第一属性值类别的第二字段;

获取所述第一字段对应的第一文字信息,以及所述第二字段对应的第二文字信息;

根据所述第一文字信息和所述第二文字信息,生成一条结构化信息。

3. 根据权利要求1所述的方法,至少一个所述字段对应的所述目标类别,为多个表头类别中的至少一个或多个表格单元类别中的至少一个;所述表头类别和所述表格单元类别之间具有一一对应关系;

其中,所述根据至少一个所述字段对应的目标类别之间的对应关系,对至少一个所述字段对应的文字信息进行组织,以得到结构化信息,包括:

根据所述目标类别为第一表头类别的第三字段,查询所述目标类别为具有对应关系的第一表格单元类别的第四字段;

获取所述第三字段对应的第三文字信息,以及所述第四字段对应的第四文字信息;

根据所述第三文字信息和所述第四文字信息,生成一条结构化信息。

4. 根据权利要求2所述的方法,其中,所述第二字段的数量为多个,所述根据所述第一文字信息和所述第二文字信息,生成一条结构化信息,包括:

对多个所述第二字段对应的第二文字信息进行合并,得到合并信息;

根据所述第一文字信息和所述合并信息,生成一条结构化信息。

5. 根据权利要求1-4任一项所述的方法,其中,所述基于至少一个所述字段对应的融合特征,确定至少一个所述字段对应的目标类别,包括:

将至少一个所述字段对应的融合特征,作为图网络中至少一个节点的节点特征,输入所述图网络中,以获取所述图网络输出的至少一个所述字段对应多个预设类别的概率;其中,至少一个所述字段与所述图网络中的节点相对应;

根据所述概率,从所述多个预设类别中,确定至少一个所述字段对应的目标类别。

6. 根据权利要求5所述的方法,所述图网络包含于结构化信息抽取模型,所述结构化信息抽取模型还包括特征提取网络,所述获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息,包括:

获取对所述待处理图像进行光学字符识别得到的所述待处理图像中至少一个所述字

段对应的文字信息,以及至少一个所述字段所属文本框的位置信息;

将所述待处理图像、所述待处理图像中至少一个所述字段对应的文字信息,以及至少一个所述字段所属文本框的位置信息,输入所述特征提取网络,以采用所述特征提取网络进行特征提取,获取所述待处理图像中至少一个所述字段对应的图像特征和文字特征。

7. 根据权利要求6所述的方法,所述特征提取网络包括卷积神经网络和循环神经网络;

其中,所述将所述待处理图像、所述待处理图像中至少一个所述字段对应的文字信息,以及至少一个所述字段所属文本框的位置信息,输入所述特征提取网络,以采用所述特征提取网络进行特征提取,获取所述待处理图像中至少一个所述字段对应的图像特征和文字特征,包括:

将所述待处理图像输入所述卷积神经网络,以采用所述卷积神经网络对所述待处理图像进行图像特征提取,得到特征图;

根据至少一个所述字段所属文本框的位置信息,截取所述特征图中对应的区域,得到至少一个所述字段对应的图像特征;

将至少一个所述字段对应的文字信息输入所述循环神经网络,以采用所述循环神经网络对所述文字信息进行文字特征提取,得到至少一个所述字段对应的文字特征。

8. 根据权利要求7所述的方法,其中,所述结构化信息抽取模型,还包括池化层,在所述根据至少一个所述字段所属文本框的位置信息,截取所述特征图中对应的区域,得到至少一个所述字段对应的图像特征之后,还包括:

将所述图像特征输入所述池化层,以采用所述池化层对所述图像特征进行池化处理,以使至少一个所述字段对应的图像特征的维度相同。

9. 根据权利要求6所述的方法,其中,所述结构化信息抽取模型,还包括特征融合层,所述将至少一个所述字段对应的所述图像特征和所述文字特征进行融合,以得到至少一个所述字段对应的融合特征,包括:

将至少一个所述字段对应的所述图像特征和所述文字特征输入所述特征融合层,以采用所述特征融合层,对至少一个所述字段对应的图像特征和所述文字特征进行拼接,或者对至少一个所述字段对应的图像特征和所述文字特征进行克罗内克积运算,得到至少一个所述字段对应的融合特征。

10. 一种用于结构化信息抽取的模型训练方法,包括:

获取训练数据及初始结构化信息抽取模型;其中,所述训练数据包括多个样本图像、多个所述样本图像中至少一个样本字段对应的样本文字信息,以及至少一个所述样本字段所属文本框的样本位置信息,至少一个所述样本字段以对应的样本类别进行标注;

将至少一个所述样本图像、所述样本图像中至少一个所述样本字段对应的样本文字信息,以及至少一个所述样本字段所属文本框的样本位置信息,输入所述初始结构化信息抽取模型,以采用所述初始结构化信息抽取模型确定至少一个所述样本字段对应的预测类别;

获取至少一个所述样本字段对应的预测类别与样本类别之间的差异,以根据所述差异对所述初始结构化信息抽取模型的模型参数进行调整,得到用于结构化信息抽取的目标结构化信息抽取模型。

11. 根据权利要求10所述的方法,其中,至少一个所述样本字段对应的所述预测类别或

样本类别,为多个属性项类别中的至少一个或多个属性值类别中的至少一个,或多个表头类别中的至少一个或多个表格单元类别中的至少一个;所述属性项类别和所述属性值类别之间具有一一对应关系;所述表头类别和所述表格单元类别之间具有一一对应关系。

12. 根据权利要求10或11所述的方法,其中,所述初始结构化信息抽取模型包括特征提取网络、特征融合层和图网络;

所述将至少一个所述样本图像、所述样本图像中至少一个所述样本字段对应的样本文字信息,以及至少一个所述样本字段所属文本框的样本位置信息,输入所述初始结构化信息抽取模型,以采用所述初始结构化信息抽取模型确定至少一个所述样本字段对应的预测类别,包括:

将所述样本图像、所述样本图像中至少一个所述样本字段对应的样本文字信息,以及至少一个所述样本字段所属文本框的样本位置信息,输入所述特征提取网络,以采用所述特征提取网络进行特征提取,获取所述样本图像中至少一个所述样本字段对应的样本图像特征和样本文字特征;

将至少一个所述样本字段对应的所述样本图像特征和所述样本文字特征输入所述特征融合层,以采用所述特征融合层,对至少一个所述样本字段对应的样本图像特征和所述样本文字特征进行融合,得到至少一个所述样本字段对应的样本融合特征;

将至少一个所述样本字段对应的样本融合特征,作为所述图网络中至少一个节点的节点特征,输入所述图网络中,以采用所述图网络对至少一个所述节点的节点特征进行特征编码,得到至少一个所述样本字段对应的预测类别。

13. 根据权利要求12所述的方法,其中,还包括:

根据所述样本图像中多个所述样本字段所属文本框的样本位置信息,获取多个所述样本字段所属文本框的中心坐标、宽度和高度;

根据多个所述样本字段所属文本框的中心坐标、宽度和高度,确定所述图网络中与多个所述样本字段对应的节点之间的边对应的初始权重参数;

根据多个所述节点之间的边对应的初始权重参数,对所述图网络中多个节点之间的边的权重矩阵进行初始化。

14. 根据权利要求13所述的方法,其中,所述根据多个所述样本字段所属文本框的中心坐标、宽度和高度,确定所述图网络中与多个所述样本字段对应的节点之间的边对应的初始权重参数,包括:

在多个所述节点中第一节点指向第二节点的情况下,根据所述第一节点和第二节点分别对应的样本字段所属文本框的中心坐标,确定所述第二节点与所述第一节点对应第一坐标轴的第一坐标差值,以及对应第二坐标轴的第二坐标差值;

对所述第一坐标差值及所述第二坐标差值作积,得到第一乘积;

对所述第一节点对应的样本字段所属文本框的宽度和高度作积,得到第二乘积;

将所述第一乘积与所述第二乘积的比值,确定为所述第一节点与所述第二节点之间的边的初始权重参数。

15. 根据权利要求12所述的方法,其中,所述将至少一个所述样本字段对应的所述样本图像特征和所述样本文字特征输入所述特征融合层,以采用所述特征融合层,对至少一个所述样本字段对应的样本图像特征和所述样本文字特征进行融合,得到至少一个所述样本

字段对应的样本融合特征,包括:

采用所述特征融合层,对至少一个所述样本字段对应的样本图像特征和所述样本文字特征进行拼接,或者对至少一个所述样本字段对应的图像特征和所述文字特征进行克罗内克积运算,得到至少一个所述样本字段对应的样本融合特征。

16. 根据权利要求12所述的方法,所述特征提取网络,包括卷积神经网络和训练神经网络;

其中,所述将所述样本图像、所述样本图像中至少一个所述样本字段对应的样本文字信息,以及至少一个所述样本字段所属文本框的样本位置信息,输入所述特征提取网络,以采用所述特征提取网络进行特征提取,获取所述样本图像中至少一个所述样本字段对应的样本图像特征和样本文字特征,包括:

将所述样本图像输入所述卷积神经网络,以采用所述卷积神经网络对所述样本图像进行图像特征提取,得到样本特征图;

根据至少一个所述样本字段所属文本框的样本位置信息,截取所述样本特征图中对应的区域,得到至少一个所述样本字段对应的样本图像特征;

将所述样本图像中至少一个所述样本字段对应的样本文字信息输入所述循环神经网络,以采用所述循环神经网络对所述样本文字信息进行文字特征提取,得到至少一个所述样本字段对应的样本文字特征。

17. 一种结构化信息抽取装置,包括:

第一获取模块,用于获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息;

融合模块,用于将至少一个所述字段对应的所述图像特征和所述文字特征进行融合,以得到至少一个所述字段对应的融合特征;

第一确定模块,用于基于至少一个所述字段对应的融合特征,确定至少一个所述字段对应的目标类别;

组织模块,用于根据至少一个所述字段对应的目标类别之间的对应关系,对至少一个所述字段对应的所述文字信息进行组织,以得到结构化信息。

18. 一种用于结构化信息抽取的模型训练装置,包括:

第二获取模块,用于获取训练数据及初始结构化信息抽取模型;其中,所述训练数据包括多个样本图像、多个所述样本图像中至少一个样本字段对应的样本文字信息,以及至少一个所述样本字段所属文本框的样本位置信息,至少一个所述样本字段以对应的样本类别进行标注;

第二确定模块,用于将至少一个所述样本图像、所述样本图像中至少一个所述样本字段对应的样本文字信息,以及至少一个所述样本字段所属文本框的样本位置信息,输入所述初始结构化信息抽取模型,以采用所述初始结构化信息抽取模型确定至少一个所述样本字段对应的预测类别;

第三获取模块,用于获取至少一个所述样本字段对应的预测类别与样本类别之间的差异,以根据所述差异对所述初始结构化信息抽取模型的模型参数进行调整,得到用于结构化信息抽取的目标结构化信息抽取模型。

19. 一种电子设备,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中任一项所述的方法,或者执行权利要求10-16中任一项所述的方法。

20.一种存储有计算机指令的非瞬时计算机可读存储介质,所述计算机指令用于使所述计算机执行权利要求1-9中任一项所述的方法,或者执行权利要求10-16中任一项所述的方法。

21.一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时实现权利要求1-9中任一项所述的方法的步骤,或者实现权利要求10-16中任一项所述的方法的步骤。

结构化信息抽取方法、装置、电子设备和存储介质

技术领域

[0001] 本公开涉及人工智能技术领域,具体为深度学习、计算机视觉技术、光学字符识别OCR技术领域,尤其涉及结构化信息抽取方法、装置、电子设备和存储介质。

背景技术

[0002] 票据识别在各行各业均有需求。例如,医疗处方是医疗报销票据中的一种,由于医疗处方数据版式繁杂且同版式间变化大,不同医院之间格式各不相同,因此,处方类数据的结构化信息抽取是图像文本识别中的一个难点。对于处方类数据等具有大量结构化文字信息的场景,如何提高结构化信息抽取的准确性,是图像文本识别中需要解决的问题。

发明内容

[0003] 本公开提供了一种用于结构化信息抽取方法、装置、电子设备和存储介质。

[0004] 根据本公开的一方面,提供了一种结构化信息抽取方法,所述方法包括:获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息;将至少一个所述字段对应的所述图像特征和所述文字特征进行融合,以得到至少一个所述字段对应的融合特征;基于至少一个所述字段对应的融合特征,确定至少一个所述字段对应的目标类别;根据至少一个所述字段对应的目标类别之间的对应关系,对至少一个所述字段对应的所述文字信息进行组织,以得到结构化信息。

[0005] 根据本公开的另一方面,提供了一种用于结构化信息抽取的模型训练方法,所述方法包括:获取训练数据及初始结构化信息抽取模型;其中,所述训练数据包括多个样本图像、多个所述样本图像中至少一个样本字段对应的样本文字信息,以及至少一个所述样本字段所属文本框的样本位置信息,至少一个所述样本字段以对应的样本类别进行标注;将至少一个所述样本图像、所述样本图像中至少一个所述样本字段对应的样本文字信息,以及至少一个所述样本字段所属文本框的样本位置信息,输入所述初始结构化信息抽取模型,以采用所述初始结构化信息抽取模型确定至少一个所述样本字段对应的预测类别;获取至少一个所述样本字段对应的预测类别与样本类别之间的差异,以根据所述差异对所述初始结构化信息抽取模型的模型参数进行调整,得到用于结构化信息抽取的目标结构化信息抽取模型。

[0006] 根据本公开的另一方面,提供了一种结构化信息抽取装置,所述装置包括:第一获取模块,用于获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息;融合模块,用于将至少一个所述字段对应的所述图像特征和所述文字特征进行融合,以得到至少一个所述字段对应的融合特征;第一确定模块,用于基于至少一个所述字段对应的融合特征,确定至少一个所述字段对应的目标类别;组织模块,用于根据至少一个所述字段对应的目标类别之间的对应关系,对至少一个所述字段对应的所述文字信息进行组织,以得到结构化信息。

[0007] 根据本公开的另一方面,提供了一种用于结构化信息抽取的模型训练装置,包括:

第二获取模块,用于获取训练数据及初始结构化信息抽取模型;其中,所述训练数据包括多个样本图像、多个所述样本图像中至少一个样本字段对应的样本文字信息,以及至少一个所述样本字段所属文本框的样本位置信息,至少一个所述样本字段以对应的样本类别进行标注;第二确定模块,用于将至少一个所述样本图像、所述样本图像中至少一个所述样本字段对应的样本文字信息,以及至少一个所述样本字段所属文本框的样本位置信息,输入所述初始结构化信息抽取模型,以采用所述初始结构化信息抽取模型确定至少一个所述样本字段对应的预测类别;第三获取模块,用于获取至少一个所述样本字段对应的预测类别与样本类别之间的差异,以根据所述差异对所述初始结构化信息抽取模型的模型参数进行调整,得到用于结构化信息抽取的目标结构化信息抽取模型。

[0008] 根据本公开的另一方面,提供了一种电子设备,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本公开的结构化信息抽取方法,或者执行本公开的用于结构化信息抽取的模型训练方法。

[0009] 根据本公开的另一方面,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,所述计算机指令用于使所述计算机执行本公开实施例公开的结构化信息抽取方法,或者执行本公开实施例公开的用于结构化信息抽取的模型训练方法。

[0010] 根据本公开的另一方面,提供了一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时实现本公开的结构化信息抽取方法的步骤,或者实现本公开的用于结构化信息抽取的模型训练方法的步骤。

[0011] 上述申请中的一个实施例具有如下优点或有益效果:

[0012] 只需确定待处理图像中至少一个字段对应的目标类别,即可根据至少一个字段对应的目标类别之间的对应关系,实现对结构化信息的抽取,从而实现了将结构化信息抽取问题简化为分类问题,简化了结构化信息抽取的过程,并且,由于结合待处理图像中至少一个字段的文字特征和图像特征两种特征进行结构化信息抽取,提高了结构化信息抽取的准确性。

[0013] 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。本公开的其它特征将通过以下的说明书而变得容易理解。

附图说明

[0014] 附图用于更好地理解本方案,不构成对本公开的限定。其中:

[0015] 图1是根据本公开第一实施例的结构化信息抽取方法的流程示意图;

[0016] 图2是根据本公开第二实施例的结构化信息抽取方法的流程示意图;

[0017] 图3是根据本公开第二实施例的结构化信息抽取模型的架构图;

[0018] 图4是根据本公开第三实施例的结构化信息抽取方法的流程示意图;

[0019] 图5是根据本公开第四实施例的用于结构化信息抽取的模型训练方法的流程示意图;

[0020] 图6是根据本公开第五实施例的用于结构化信息抽取的模型训练方法的流程示意图;

[0021] 图7是根据本公开第六实施例的结构化信息抽取装置的结构示意图

- [0022] 图8是根据本公开第七实施例的结构化信息抽取装置的结构示意图；
- [0023] 图9是根据本公开第八实施例的用于结构化信息抽取的模型训练装置的结构示意图；
- [0024] 图10是根据本公开第九实施例的用于结构化信息抽取的模型训练装置的结构示意图；
- [0025] 图11是用来实现本公开实施例的结构化信息抽取方法或用于结构化信息抽取的模型训练方法的电子设备的框图。

具体实施方式

[0026] 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。

[0027] 本公开提供的结构化信息抽取方法、用于结构化信息抽取的模型训练方法、装置、电子设备、非瞬时计算机可读存储介质以及计算机程序产品,涉及人工智能技术领域,具体为深度学习、计算机视觉技术、光学字符识别OCR技术领域。

[0028] 其中,人工智能是研究使计算机来模拟人的某些思维过程和智能行为(如学习、推理、思考、规划等)的学科,既有硬件层面的技术也有软件层面的技术。人工智能硬件技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理等技术;人工智能软件技术主要包括计算机视觉、语音识别技术、自然语言处理技术以及机器学习/深度学习、大数据处理技术、知识图谱技术等几大方向。

[0029] 下面参考附图描述本公开实施例的结构化信息抽取方法、用于结构化信息抽取的模型训练方法、装置、电子设备、非瞬时计算机可读存储介质以及计算机程序产品。

[0030] 图1是根据本公开第一实施例的结构化信息抽取方法的流程示意图。其中,需要说明的是,本实施的结构化信息抽取方法,执行主体为结构化信息抽取装置,该结构化信息抽取装置可以由软件和/或硬件实现,该结构化信息抽取装置可以配置在电子设备,该电子设备可以包括但不限于终端设备、服务器等,该实施例对电子设备不作具体限定。

[0031] 如图1所示,该结构化信息抽取方法可以包括:

[0032] 步骤101,获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息。

[0033] 其中,待处理图像,可以是通过实体文本进行拍摄、扫描等方式得到的、具有结构化文字信息的图像,比如医疗处方类的票据图像、表单图像、证件图像等,本公开对此不作限制。

[0034] 在本公开的实施例中,结构化信息抽取装置,可以获取待处理图像中每个字段分别对应的图像特征、文字特征以及文字信息。

[0035] 其中,字段,由连续的字组成,比如票据图像中的“姓名”为一个字段,“性别”为一个字段,“药品名称”为一个字段,某个具体的人名“张三”为一个字段。

[0036] 其中,图像特征,用于表征字段的图像视觉信息。文字特征,基于字段的文字内容得到,用于表征字段的文字含义,比如可以包括字段的语义特征。文字信息,为字段对应的

文字内容信息。

[0037] 步骤102,将至少一个字段对应的图像特征和文字特征进行融合,以得到至少一个字段对应的融合特征。

[0038] 其中,字段对应的融合特征,同时包含字段对应的图像特征和文字特征,比如可以通过将字段对应的图像特征和文字特征进行拼接,得到字段对应的融合特征。

[0039] 步骤103,基于至少一个字段对应的融合特征,确定至少一个字段对应的目标类别。

[0040] 在本公开的实施例中,可以预先设置多个类别,并且预设多个类别之间具有一一对应关系。

[0041] 比如,在相关技术中,通常将具有key-value对应关系的字段划分为属性项key和属性值value两种类别。而本公开实施例中,对于属性项,可以进一步划分为多个属性项类别,比如,作为属性项的“年龄”、“性别”和“地址”,可以进一步划分为三个属性项类别:年龄项类别、性别项类别、地址项类别;类似的,对于属性值,也可以进一步划分为多个属性值类别,比如,将属性项“年龄”、“性别”和“地址”分别对应的属性值,进一步划分为三个属性值类别:年龄值类别、性别值类别、地址值类别。其中,划分后的年龄项类别与年龄值类别具有对应关系,性别项类别与性别值类别具有对应关系,地址项类别与地址值类别具有对应关系。

[0042] 进而,可以基于至少一个字段对应的融合特征进行推理,从预设的多个类别中,确定各个字段分别对应的目标类别。比如,每个字段对应的目标类别,可以为上述多个属性项类别中的一个,或上述多个属性值类别中的一个。

[0043] 步骤104,根据至少一个字段对应的目标类别之间的对应关系,对至少一个字段对应的文字信息进行组织,以得到结构化信息。

[0044] 在本公开的实施例中,由于预设的多个类别之间具有一一对应关系,从而在确定各字段对应的目标类别之后,可以根据各字段对应的目标类别之间的对应关系,将对应的目标类别之间具有对应关系的各字段的文字信息进行组合,从而得到一条或多条结构化信息。

[0045] 举例来说,假设对于某张待处理图像,获取了其中包括的4字段分别对应的图像特征、文字特征和文字信息,其中,这4个字段分别对应的文字信息为“姓名”、“年龄”、“18”、“张三”。通过对这4个字段分别对应的图像特征和文字特征进行融合,得到了这4个字段分别对应的融合特征,进而基于这4个字段分别对应的融合特征,确定了这4个字段分别对应的目标类别。其中,文字信息为“姓名”的字段1对应的是姓名项类别;文字信息为“年龄”的字段2对应的是年龄项类别;文字信息为“18”的字段3对应的是年龄值类别;文字信息为“张三”的字段4对应的是姓名值类别。

[0046] 由于字段2对应的年龄项类别,和字段3对应的年龄值类别具有对应关系,字段1对应的姓名项类别,和字段4对应的姓名值类别具有对应关系,则可以将字段1对应的“姓名”,与字段4对应的“张三”组合,得到一条结构化信息,将字段2对应的“年龄”,与字段3对应的“18”组合,得到一条结构化信息。

[0047] 通过预设多个类别,并且预设多个类别之间具有一一对应关系,从而在从预设的多个类别中确定各字段分别对应的目标类别后,可以仅基于各字段分别对应的目标类别之

间的对应关系,对各字段分别对应的文字信息进行组织,得到结构化信息。并且,由于结合待处理图像中至少一个字段的文字特征和图像特征两种特征进行结构化信息抽取,能够提高结构化信息抽取的准确性,从而对于自然场景中存在错行、版式复杂等问题的医疗明细小票类票据等数据,能够实现结构化信息的准确抽取。

[0048] 本公开实施例的结构化信息抽取方法,通过获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息,将至少一个字段对应的图像特征和文字特征进行融合,以得到至少一个字段对应的融合特征,基于至少一个字段对应的融合特征,确定至少一个字段对应的目标类别,根据至少一个字段对应的目标类别之间的对应关系,对至少一个字段对应的文字信息进行组织,以得到结构化信息,由此,只需确定待处理图像中至少一个字段对应的目标类别,即可根据至少一个字段对应的目标类别之间的对应关系,实现对结构化信息的抽取,从而实现了将结构化信息抽取问题简化为分类问题,简化了结构化信息抽取的过程,并且,由于结合待处理图像中至少一个字段的文字特征和图像特征两种特征进行结构化信息抽取,提高了结构化信息抽取的准确性。

[0049] 下面结合图2,对本公开提供的结构化信息抽取方法中,获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息的过程,对至少一个字段对应的图像特征和文字特征进行融合得到融合特征的过程,以及基于至少一个字段对应的融合特征,确定至少一个字段对应的目标类别的过程进一步说明。

[0050] 图2是根据本公开第二实施例的结构化信息抽取方法的流程示意图。如图2所示,结构化信息抽取方法,可以包括以下步骤:

[0051] 步骤201,获取对待处理图像进行光学字符识别得到的待处理图像中至少一个字段对应的文字信息,以及至少一个字段所属文本框的位置信息。

[0052] 其中,字段所属文本框的位置信息,具体可以为字段所属文本框的四个角点各自的坐标。

[0053] 在本公开实施例中,可以将待处理图像输入光学字符识别接口,对待处理图像进行光学字符识别,以获取待处理图像中各字段对应的文字信息、各字段所属文本框的位置信息。结构化信息抽取装置可以与光学字符识别接口对接,从而可以通过光学字符识别接口,获取待处理图像中各字段对应的文字信息、各字段所属文本框的位置信息。

[0054] 步骤202,将待处理图像、待处理图像中至少一个字段对应的文字信息,以及至少一个字段所属文本框的位置信息,输入特征提取网络,以采用特征提取网络进行特征提取,获取待处理图像中至少一个字段对应的图像特征和文字特征。

[0055] 在本公开实施例中,可以预先训练一个端到端的结构化信息抽取模型,结构化信息抽取模型的输入为图像、该图像中至少一个字段对应的文字信息,以及该图像中至少一个字段所属文本框的位置信息,输出为该图像中至少一个字段对应的类别。具体的训练过程,可以参考下述实施例的用于结构化信息抽取的模型训练方法的实施例的描述,此处不再赘述。

[0056] 其中,结构化信息抽取模型中包括特征提取网络,用于提取图像中至少一个字段对应的图像特征和文字特征,从而可以将步骤201中获取的待处理图像、待处理图像中至少一个字段对应的文字信息,以及至少一个字段所属文本框的位置信息,输入特征提取网络,以采用特征提取网络进行特征提取,获取待处理图像中至少一个字段对应的图像特征和文

字特征。

[0057] 通过采用特征提取网络,基于待处理图像及其中至少一个字段对应的文字信息,以及至少一个字段所属文本框的位置信息,提取待处理图像中至少一个字段对应的图像特征和文字特征,能够实现准确的图像特征和文字特征的提取,为后续基于图像特征和文字特征实现准确的结构化信息抽取奠定基础。

[0058] 在本公开的实施例中,参考图3,特征提取网络,可以包括卷积神经网络301和循环神经网络302。其中,卷积神经网络301,可以为任意类型的能够实现图像特征提取的卷积神经网络,本公开对此不作限制。循环神经网络302,可以为任意类型的能够实现文字特征提取的循环神经网络,本公开对此不作限制。

[0059] 在本公开实施例中,参考图3,可以采用如下方式,获取待处理图像中至少一个字段对应的图像特征和文字特征:

[0060] 将待处理图像输入卷积神经网络301,以采用卷积神经网络301对待处理图像进行图像特征提取,得到特征图;根据至少一个字段所属文本框的位置信息,截取特征图中对应的区域,得到至少一个字段对应的图像特征;将至少一个字段对应的文字信息输入循环神经网络302,以采用循环神经网络302对文字信息进行文字特征提取,得到至少一个字段对应的文字特征。

[0061] 其中,可以由结构化信息抽取模型的其中一个处理层,实现根据至少一个字段所属文本框的位置信息,截取特征图中对应的区域,得到至少一个字段对应的图像特征。

[0062] 通过上述过程,实现了基于卷积神经网络和循环神经网络,准确提取待处理图像中至少一个字段对应的图像特征和文字特征,为后续基于图像特征和文字特征实现准确的结构化信息抽取奠定基础。

[0063] 可以理解的是,采用卷积神经网络对待处理图像进行图像特征提取,得到的是与待处理图像的大小相同的特征图,而由于一张图像中各字段包括的字数可能是不同的,相应的,各字段所属文本框的大小可能不同,从而根据各字段所属文本框的位置信息,在特征图中截取的文本框对应区域的大小也可能不同,进而各字段对应的图像特征的维度可能不同。其中,维度包括宽度、高度和通道数。而对于一个模型来说,要求输入的各特征的维度相同。那么,本公开实施例中,根据至少一个字段所属文本框的位置信息,截取特征图中对应的区域,得到至少一个字段对应的图像特征后,还可以对至少一个字段对应的图像特征的维度进行处理,以使各字段对应的图像特征的维度相同。

[0064] 相应的,结构化信息抽取模型,还可以包括池化层(图3中未示出),从而在根据至少一个字段所属文本框的位置信息,截取特征图中对应的区域,得到至少一个字段对应的图像特征之后,还可以将图像特征输入池化层,以采用池化层对至少一个字段对应的图像特征进行池化处理,以使至少一个字段对应的图像特征的维度相同,即使至少一个字段对应的图像特征的宽度、高度和通道数相同。

[0065] 步骤203,将至少一个字段对应的图像特征和文字特征输入特征融合层,以采用特征融合层,对至少一个字段对应的图像特征和文字特征进行拼接,或者对至少一个字段对应的图像特征和文字特征进行克罗内克积运算,得到至少一个字段对应的融合特征。

[0066] 在本公开的实施例中,参考图3,结构化信息抽取模型中还可以包括特征融合层303,用于对至少一个字段对应的图像特征和文字特征进行融合,从而可以将至少一个字段

对应的图像特征和文字特征输入特征融合层,以采用特征融合层,对至少一个字段对应的图像特征和文字特征进行融合,得到至少一个字段对应的融合特征。

[0067] 在本公开的实施例中,针对每个字段,可以采用多种方式,实现对对应的图像特征和文字特征的融合。作为一个可能的实现方式,针对每个字段,可以将字段对应的图像特征和文字特征进行拼接,将拼接后的拼接特征作为字段对应的融合特征。或者,作为另一种可能的实现方式,针对每个字段,可以将字段对应的图像特征和文字特征的克罗内克积,作为字段的融合特征。通过采用特征融合层,对至少一个字段对应的图像特征和文字特征进行融合,实现了对于待处理图像中的各字段,获取各字段对应的融合特征,且通过采用克罗内克积运算,将字段对应的图像特征和文字特征进行融合,可以实现较好的特征融合效果。

[0068] 需要说明的是,上述特征融合方式仅是示例性说明,不能理解为对本公开技术方案的限制,在实际应用中,本领域技术人员可以采用其它任意方式实现对字段对应的图像特征和文字特征的融合,本公开对此不作限制。

[0069] 步骤204,将至少一个字段对应的融合特征,作为图网络中至少一个节点的节点特征,输入图网络中,以获取图网络输出的至少一个字段对应多个预设类别的概率;其中,至少一个字段与图网络中的节点相对应。

[0070] 在本公开的实施例中,参考图3,结构化信息抽取模型,还可以包括图网络304,用于基于图像中至少一个字段对应的融合特征,确定至少一个字段对应的类别,从而可以将待处理图像中至少一个字段对应的融合特征,作为图网络中至少一个节点的节点特征,输入图网络中,以采用图网络对至少一个节点的节点特征进行特征编码,得到至少一个字段对应的目标类别。其中,每个字段对应图网络中的一个节点。

[0071] 在本公开的实施例中,可以预先设置多个类别(即步骤204中的多个预设类别),获取待处理图像中至少一个字段对应的融合特征后,可以将至少一个字段对应的融合特征,作为图网络中至少一个节点的节点特征,输入图网络中,以获取图网络输出的至少一个字段对应多个预设类别的概率。即可以采用图网络,预测每个字段分别对应多个预设类别的概率。

[0072] 步骤205,根据概率,从多个预设类别中,确定至少一个字段对应的目标类别。

[0073] 在本公开实施例中,针对每个字段,可以将多个预设类别中,对应的概率最大的预设类别,确定为该字段对应的目标类别。

[0074] 通过上述过程,实现了采用图网络,基于待处理图像中各字段对应的融合特征,准确确定各字段对应的目标类别,为后续根据各字段对应的目标类别之间的对应关系,进行准确的结构化信息抽取奠定了基础。

[0075] 步骤206,根据至少一个字段对应的目标类别之间的对应关系,对至少一个字段对应的文字信息进行组织,以得到结构化信息。

[0076] 其中,步骤206的具体实现过程及原理,可以参考上述实施例的描述,此处不再赘述。

[0077] 本公开实施例的结构化信息抽取方法,实现了采用端到端模型,基于待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息,将至少一个字段对应的图像特征和文字特征进行融合,以得到至少一个字段对应的融合特征,并基于至少一个字段对应的融合特征,确定至少一个字段对应的目标类别,进而使得能够仅根据至少一个字段对应

的目标类别之间的对应关系,对至少一个字段对应的文字信息进行组织,以得到结构化信息,由此,实现了将结构化信息抽取问题简化为分类问题,简化了结构化信息抽取的过程,并且,由于结合待处理图像中至少一个字段的文字特征和图像特征两种特征进行结构化信息抽取,提高了结构化信息抽取的准确性。

[0078] 下面结合图4,对本公开提供的结构化信息抽取方法中,根据至少一个字段对应的目标类别之间的对应关系,对至少一个字段对应的文字信息进行组织,得到结构化信息的过程进一步说明。

[0079] 图4是根据本公开第三实施例的结构化信息抽取方法的流程示意图。如图4所示,结构化信息抽取方法,可以包括以下步骤:

[0080] 步骤401,获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息。

[0081] 步骤402,将至少一个字段对应的图像特征和文字特征进行融合,以得到至少一个字段对应的融合特征。

[0082] 其中,步骤401-402的具体实现过程及原理,可以参考上述实施例的描述,此处不再赘述。

[0083] 步骤403,基于至少一个字段对应的融合特征,确定至少一个字段对应的目标类别。

[0084] 在本公开实施例中,可以预先训练一个图网络,从而可以将至少一个字段对应的融合特征作为图网络中至少一个节点的节点特征,输入图网络,以获取图网络输出的至少一个字段对应多个预设类别的概率,其中,至少一个字段与图网络中的节点相对应,进而针对每个字段,根据上述概率,将多个预设类别中,对应的概率最大的预设类别,确定为该字段对应的目标类别。

[0085] 通过上述过程,实现了采用图网络,基于待处理图像中各字段对应的融合特征,准确确定各字段对应的目标类别,为后续根据各字段对应的目标类别之间的对应关系,进行准确的结构化信息抽取奠定了基础。

[0086] 在本公开的实施例中,可以预先设置多个类别(即前述实施例中的多个预设类别),并且预设多个类别之间具有一一对应关系,进而,可以基于至少一个字段对应的融合特征进行推理,从预设的多个类别中,确定各个字段分别对应的目标类别,进而根据各字段对应的目标类别之间的对应关系,将对应的目标类别之间具有对应关系的字段的文字信息进行组合,从而得到一条或多条结构化信息。

[0087] 可以理解的是,在相关技术中,通常将具有key-value对应关系的字段划分为属性项key和属性值value两种类别。而本公开实施例中,对于属性项,可以进一步划分为多个属性项类别,比如,作为属性项的“年龄”、“性别”和“地址”,可以进一步划分为三个属性项类别:年龄项类别、性别项类别、地址项类别;类似的,对于属性值,也可以进一步划分为多个属性值类别,比如,将属性项“年龄”、“性别”和“地址”分别对应的属性值,进一步划分为三个属性值类别:年龄值类别、性别值类别、地址值类别。其中,划分后的年龄项类别与年龄值类别具有对应关系,性别项类别与性别值类别具有对应关系,地址项类别与地址值类别具有对应关系。

[0088] 即,在本公开的实施例中,可以预先设置多个属性项类别、多个属性值类别,多个

属性项类别与多个属性值类别之间具有一一对应关系,待处理图像中至少一个字段对应的目标类别,为多个属性项类别中的至少一个或多个属性值类别中的至少一个。相应的,可以通过以下步骤404-406的方式,实现根据待处理图像中至少一个字段对应的目标类别之间的对应关系,对至少一个字段对应的文字信息进行组织,以得到结构化信息。

[0089] 步骤404,根据目标类别为第一属性项类别的第一字段,查询目标类别为具有对应关系的第一属性值类别的第二字段。

[0090] 其中,第一属性项类别,为多个属性项类别中的一个。第一属性值类别,为多个属性值类别中的一个。第一字段和第二字段,分别为待处理图像中至少一个字段中的其中一个字段。

[0091] 举例来说,假设待处理图像中的某个第一字段对应的第一属性项类别为姓名项类别,由于与姓名项类别具有对应关系的第一属性值类别为姓名值类别,则可以从待处理图像中的各字段中,查询目标类别为姓名值类别的第二字段。

[0092] 步骤405,获取第一字段对应的第一文字信息,以及第二字段对应的第二文字信息。

[0093] 其中,获取第一字段对应的第一文字信息的步骤,也可以在查询目标类别为具有对应关系的第一属性值类别的第二字段之前,本公开实施例对获取第一字段对应的第一文字信息的步骤的执行时机不作限制。

[0094] 步骤406,根据第一文字信息和第二文字信息,生成一条结构化信息。

[0095] 继续上述示例,查询到目标类别为姓名值类别的第二字段后,可以获取第二字段对应的第二文字信息,继而将第一文字信息和第二文字信息进行组合,生成一条结构化信息。

[0096] 通过将属性项类别和属性值类别划分为更细粒度的类别,并预设属性项类别与属性值类别之间具有一一对应关系,进而根据目标类别为第一属性项类别的第一字段,查询目标类别为具有对应关系的第一属性值类别的第二字段,获取第一字段对应的第一文字信息,以及第二字段对应的第二文字信息,根据第一文字信息和第二文字信息,生成一条结构化信息,实现了仅根据至少一个字段对应的目标类别之间的对应关系,实现对待处理图像中目标类别为属性项类别和属性值类别的字段的结构化信息的抽取。

[0097] 可以理解的是,在实际应用中,对于目标类别为某个属性项类别的某个字段,目标类别为与该属性项类别具有对应关系的属性值类别的字段可能为多个,即上述第二字段的数量可能为多个,在这种情况下,可以对多个第二字段对应的第二文字信息进行合并,得到合并信息,进而根据第一文字信息和合并信息,生成一条结构化信息。

[0098] 举例来说,假设待处理图像中某个地点的地址包括3行,其中每行为一个字段,这3个字段对应的目标类别均为地址值类别,与地址项类别对应。相应的,根据目标类别为地址项类别的第一字段,可以查询到目标类别为地址值类别的第二字段的数量为3个,从而可以将这3个字段分别对应的文字信息进行合并,即合并为一条地址信息,进而可以将该合并的地址信息,与目标类别为地址项类别的第一字段的第一文字信息进行组合,生成一条结构化信息。

[0099] 通过在第二字段的数量为多个时,对多个第二字段对应的第二文字信息进行合并,得到合并信息,进而根据第一文字信息和合并信息,生成一条结构化信息,实现了在某

个属性项类别与某个属性值类别具有对应关系,且目标类别为该属性项类别的第一字段和目标类别为该属性值类别的第二字段的数量为一对多的情况下,仅根据各字段对应的目标类别之间的对应关系,对各字段对应的文字信息进行组织,得到准确的结构化信息。

[0100] 在本公开的实施例中,与上述对多个属性项类别和多个属性值类别的划分方式类似,对于具有对应关系的表格的表头和表格的单元这两种类别,也可以分别进一步划分为多个表头类别和多个表格单元类别。其中,划分后的多个表头类别和多个表格单元类别之间具有一一对应关系。至少一个字段对应的目标类别,为多个表头类别中的至少一个或多个表格单元类别中的至少一个。

[0101] 相应的,可以通过以下方式,实现根据至少一个字段对应的目标类别之间的对应关系,对至少一个字段对应的文字信息进行组织,以得到结构化信息:根据目标类别为第一表头类别的第三字段,查询目标类别为具有对应关系的第一表格单元类别的第四字段;获取第三字段对应的第三文字信息,以及第四字段对应的第四文字信息;根据第三文字信息和第四文字信息,生成一条结构化信息。

[0102] 其中,第一表头类别,为多个表头类别中的一个。第一表格单元类别,为多个表格单元类别中的一个。第三字段和第四字段,分别为待处理图像中至少一个字段中的其中一个字段。

[0103] 举例来说,假设待处理图像中的某个第三字段对应的第一表头类别为A类别,第三字段对应的是第三文字信息,由于与A类别具有对应关系的第一表格单元类别为B类别,则可以从待处理图像中的各字段中,查询目标类别为B类别的第四字段。查询到目标类别为B类别的第四字段后,可以获取第四字段对应的第四文字信息,继而将第三文字信息和第四文字信息进行组合,生成一条结构化信息。

[0104] 通过将表头类别和表格单元类别划分为更细粒度的类别,并预设表头类别与表格单元类别之间具有一一对应关系,进而根据目标类别为第一表头类别的第三字段,查询目标类别为具有对应关系的第一表格单元类别的第四字段,获取第三字段对应的第三文字信息,以及第四字段对应的第四文字信息,根据第三文字信息和第四文字信息,生成一条结构化信息,实现了仅根据至少一个字段对应的目标类别之间的对应关系,实现对待处理图像中目标类别为表头类别和表格单元类别的字段的结构化信息的抽取。

[0105] 需要说明的是,在一种可能的实现形式中,结构化信息的抽取结果,不仅可以包括目标类别为属性项类别的第一字段对应的第一识别信息和目标类别为属性值类别的第二字段对应的第二识别信息,或者目标类别为表头类别的第三字段对应的第三识别信息和目标类别为表格单元的第四字段对应的第四识别信息,还可以包括确定的第一字段、第二字段、第三字段或第四字段为对应的目标类别的概率、字段所属文本框的位置信息等信息,本公开对此不作限制。

[0106] 根据本公开的实施例,还提供一种用于结构化信息抽取的模型训练方法。

[0107] 图5是根据本公开第四实施例的用于结构化信息抽取的模型训练方法的流程示意图。

[0108] 其中,需要说明的是,本公开实施例提供的用于结构化信息抽取的模型训练方法,执行主体为用于结构化信息抽取的模型训练装置,以下简称模型训练装置。该模型训练装置可以由软件和/或硬件实现,该模型训练装置可以配置在电子设备,该电子设备可以包括

但不限于终端设备、服务器等,该实施例对电子设备不作具体限定。

[0109] 如图5所示,用于结构化信息抽取的模型训练方法,可以包括以下步骤:

[0110] 步骤501,获取训练数据及初始结构化信息抽取模型。

[0111] 其中,训练数据包括多个样本图像、多个样本图像中至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,至少一个样本字段以对应的样本类别进行标注。

[0112] 其中,初始结构化信息抽取模型,为一个端到端的模型,初始结构化信息抽取模型的输入为图像、该图像中至少一个字段对应的文字信息,以及该图像中至少一个字段所属文本框的位置信息,输出为该图像中至少一个字段对应的类别。

[0113] 在本公开实施例中,训练数据中包括的每个样本图像中的至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,可以采用光学字符识别技术提前获取,从而在进行模型训练时,可以直接作为已知的光学字符识别结果用于进行模型训练。

[0114] 步骤502,将至少一个样本图像、样本图像中至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,输入初始结构化信息抽取模型,以采用初始结构化信息抽取模型确定至少一个样本字段对应的预测类别。

[0115] 在本公开的实施例中,初始结构化信息抽取模型可以基于样本图像、样本图像中至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,对至少一个样本字段对应的类别进行预测,得到至少一个样本字段对应的预测类别。

[0116] 步骤503,获取至少一个样本字段对应的预测类别与样本类别之间的差异,以根据差异对初始结构化信息抽取模型的模型参数进行调整,得到用于结构化信息抽取的目标结构化信息抽取模型。

[0117] 在本公开的实施例中,得到至少一个样本字段对应的预测类别后,即可获取至少一个样本字段对应的预测类别与样本类别之间的差异,进而根据差异对初始结构化信息抽取模型的模型参数进行调整,得到用于结构化信息抽取的目标结构化信息抽取模型。

[0118] 其中,对初始结构化信息抽取模型进行训练时,初始结构化信息抽取模型的损失函数可以为以下公式(1)所示的样本字段的类别的交叉熵损失。

[0119]
$$L = -t_j \log(y_j) \quad (1)$$

[0120] 其中,L为每个样本字段对应的类别的交叉熵损失。 t_j 为样本字段对应的样本类别, y_j 为样本字段对应的预测类别。 j 表示第 j 类。

[0121] 其中,对初始结构化信息抽取模型进行模型训练时,可以采用深度学习的方法,相比其它训练方法,深度学习方法在大数据上的表现效果更好。其中,根据至少一个样本字段对应的预测类别与样本类别之间的差异,对初始结构化信息抽取模型进行训练的过程,可以参考相关技术,本公开对此不作赘述。

[0122] 需要说明的是,本公开实施例中的训练后得到的目标结构化信息抽取模型,可以用于执行以下步骤:

[0123] 获取待处理图像中至少一个字段对应的图像特征和文字特征;

[0124] 将至少一个字段对应的图像特征和文字特征进行融合,以得到至少一个字段对应的融合特征;

[0125] 基于至少一个字段对应的融合特征,确定至少一个字段对应的目标类别,其中,至少一个字段对应的目标类别之间的对应关系,用于对至少一个字段对应的文字信息进行组织,以得到结构化信息。

[0126] 采用训练后得到的目标结构化信息抽取模型执行上述步骤的过程,可以参见上述结构化信息抽取方法的实施例的说明,此处不再赘述。

[0127] 综上,本公开实施例提供的用于结构化信息抽取的模型训练方法,通过获取训练数据及初始结构化信息抽取模型,将至少一个样本图像、样本图像中至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,输入初始结构化信息抽取模型,以采用初始结构化信息抽取模型确定至少一个样本字段对应的预测类别,获取至少一个样本字段对应的预测类别与样本类别之间的差异,以根据差异对初始结构化信息抽取模型的模型参数进行调整,得到用于结构化信息抽取的目标结构化信息抽取模型,实现了基于训练数据,对初始结构化信息抽取模型进行模型训练,得到用于结构化信息抽取的目标结构化信息抽取模型,利用训练后得到的目标结构化信息抽取模型进行结构化信息抽取,能够实现将结构化信息抽取问题简化为分类问题,简化结构化信息抽取的过程,并且,通过结合待处理图像中至少一个字段的文字特征和图像特征两种特征进行结构化信息抽取,提高结构化信息抽取的准确性。

[0128] 下面结合图6,对本公开提供的用于结构化信息抽取的模型训练装置进一步说明。

[0129] 图6是根据本公开第五实施例的用于结构化信息抽取的模型训练方法的流程示意图。

[0130] 如图6所示,用于结构化信息抽取的模型训练方法,可以包括以下步骤:

[0131] 步骤601,获取训练数据及初始结构化信息抽取模型。

[0132] 其中,训练数据包括多个样本图像、多个样本图像中至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,至少一个样本字段以对应的样本类别进行标注。

[0133] 其中,步骤601的具体实现过程及原理,可以参考上述实施例的描述,此处不再赘述。

[0134] 步骤602,将样本图像、样本图像中至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,输入特征提取网络,以采用特征提取网络进行特征提取,获取样本图像中至少一个样本字段对应的样本图像特征和样本文字特征。

[0135] 在本公开的实施例中,初始结构化信息抽取模型可以包括特征提取网络,用于提取图像中至少一个字段对应的图像特征和文字特征,从而可以将样本图像、样本图像中至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,输入特征提取网络,以采用特征提取网络进行特征提取,获取样本图像中至少一个样本字段对应的样本图像特征和样本文字特征。

[0136] 在本公开实施例中,特征提取网络,可以包括卷积神经网络和训练神经网络,相应的,步骤602可以通过以下方式实现:

[0137] 将样本图像输入卷积神经网络,以采用卷积神经网络对样本图像进行图像特征提取,得到样本特征图;

[0138] 根据至少一个样本字段所属文本框的样本位置信息,截取样本特征图中对应的区

域,得到至少一个样本字段对应的样本图像特征;

[0139] 将样本图像中至少一个样本字段对应的样本文字信息输入循环神经网络,以采用循环神经网络对样本文字信息进行文字特征提取,得到至少一个样本字段对应的样本文字特征。

[0140] 由此,实现了基于卷积神经网络和循环神经网络,准确提取样本图像中至少一个样本字段对应的样本图像特征和样本文字特征,为后续基于样本图像特征和样本文字特征,准确预测至少一个样本字段对应的预测类别,从而提高训练后得到的目标结构化信息抽取模型的性能。

[0141] 其中,通过卷积神经网络和循环神经网络,提取样本图像中至少一个样本字段对应的样本图像特征和样本文字特征的过程,可以参考基于卷积神经网络和循环神经网络,提取待处理图像中至少一个字段对应的图像特征和文字特征的过程,此处不再赘述。

[0142] 步骤603,将至少一个样本字段对应的样本图像特征和样本文字特征输入特征融合层,以采用特征融合层,对至少一个样本字段对应的样本图像特征和样本文字特征进行融合,得到至少一个样本字段对应的样本融合特征。

[0143] 在本公开的实施例中,初始结构化信息抽取模型中还可以包括特征融合层,用于对至少一个字段对应的图像特征和文字特征进行融合,从而可以将至少一个样本字段对应的样本图像特征和样本文字特征输入特征融合层,以采用特征融合层,对至少一个样本字段对应的样本图像特征和样本文字特征进行融合,得到至少一个样本字段对应的样本融合特征。

[0144] 在本公开的实施例中,针对每个样本字段,可以采用多种方式,实现对对应的样本图像特征和样本文字特征的融合。作为一个可能的实现方式,针对每个样本字段,可以将样本字段对应的样本图像特征和样本文字特征进行拼接,将拼接后的拼接特征作为样本字段对应的样本融合特征。或者,作为另一种可能的实现方式,针对每个样本字段,可以将样本字段对应的样本图像特征和样本文字特征的克罗内克积,作为样本字段的样本融合特征。通过采用特征融合层,对至少一个样本字段对应的样本图像特征和样本文字特征进行融合,实现了对于样本图像中的各样本字段,获取各样本字段对应的样本融合特征,且通过采用克罗内克积运算,将样本字段对应的样本图像特征和样本文字特征进行融合,可以实现较好的特征融合效果。

[0145] 需要说明的是,上述特征融合方式仅是示例性说明,不能理解为对本公开技术方案的限制,在实际应用中,本领域技术人员可以采用其它任意方式实现对样本图像特征和样本文字特征的融合,本公开对此不作限制。

[0146] 步骤604,将至少一个样本字段对应的样本融合特征,作为图网络中至少一个节点的节点特征,输入图网络中,以采用图网络对至少一个节点的节点特征进行特征编码,得到至少一个样本字段对应的预测类别。

[0147] 在本公开的实施例中,初始结构化信息抽取模型还可以包括图网络,用于基于图像中至少一个字段对应的融合特征,确定该图像中至少一个字段对应的类别,从而可以将该图像中至少一个样本字段对应的样本融合特征,作为图网络中至少一个节点的节点特征,输入图网络中,以采用图网络对至少一个节点的节点特征进行特征编码,得到至少一个样本字段对应的预测类别。

[0148] 在本公开的实施例中,可以预先设置多个类别,获取样本图像中至少一个样本字段对应的样本融合特征后,可以将至少一个样本字段对应的样本融合特征,作为图网络中至少一个节点的节点特征,输入图网络中,以获取图网络输出的至少一个样本字段对应多个预设类别的概率。即可以采用图网络,预测每个样本字段分别对应多个预设类别的概率。进而针对每个样本字段,可以将多个预设类别中,对应的概率最大的预设类别,确定为该样本字段对应的预测类别。

[0149] 步骤605,获取至少一个样本字段对应的预测类别与样本类别之间的差异,以根据差异对初始结构化信息抽取模型的模型参数进行调整,得到用于结构化信息抽取的目标结构化信息抽取模型。

[0150] 在本公开的实施例中,训练后得到的目标结构化信息抽取模型用于执行以下步骤:获取待处理图像中至少一个字段对应的图像特征和文字特征;将至少一个字段对应的图像特征和文字特征进行融合,以得到至少一个字段对应的融合特征;基于至少一个字段对应的融合特征,确定至少一个字段对应的目标类别,其中,至少一个字段对应的目标类别,用于对至少一个字段对应的文字信息进行组织,以得到结构化信息。

[0151] 其中,至少一个样本字段对应的预测类别、样本类别或目标类别,为多个属性项类别中的至少一个或多个属性值类别中的至少一个,或多个表头类别中的至少一个或多个表格单元类别中的至少一个;属性项类别和属性值类别之间具有一一对应关系;表头类别和表格单元类别之间具有一一对应关系。通过将属性项类别、属性值类别、表头类别和表格单元类别分别划分为更细粒度的类别,使得训练得到的目标结构化信息抽取模型能够进行更细粒度的类别预测,进而实现仅根据待处理图像中至少一个字段之间的对应关系,进行结构化信息的抽取,从而简化结构化信息的抽取过程。

[0152] 通过端到端的初始结构化信息抽取模型包括特征提取网络、特征融合层和图网络,采用特征提取网络进行样本图像特征和样本文字特征的提取,采用特征融合层进行样本图像特征和样本文字特征的融合,采用图网络确定样本字段对应的预测类别,进而根据各样本字段对应的预测类别与样本类别之间的差异,对初始结构化信息抽取模型的模型参数进行调整,得到用于结构化信息抽取的目标结构化信息抽取模型,使得无需单独分别对特征提取网络、特征融合层、图网络进行训练,通过一次训练即可得到训练好的特征提取网络、特征融合层和图网络,从而能够减少特征提取网络、特征融合层和图网络的训练成本。

[0153] 在本公开的实施例中,在获取初始结构化信息抽取模型之后,还可以通过图网络中各节点的空间布局信息,比如两个节点之间的距离,对图网络中节点与节点之间的边的权重矩阵进行初始化,从而在对初始结构化信息抽取模型进行训练时,可以根据至少一个样本字段对应的预测类别与样本类别之间的差异,对初始化后的初始结构化信息抽取模型的模型参数进行调整,从而提高后续训练得到的目标结构化信息抽取模型的模型性能,提高分类准确性。

[0154] 具体的,可以采用如下方式,对图网络中至少一个节点之间的边的权重矩阵进行初始化:根据样本图像中多个样本字段所属文本框的样本位置信息,获取多个样本字段所属文本框的中心坐标、宽度和高度;根据多个样本字段所属文本框的中心坐标、宽度和高度,确定图网络中与多个样本字段对应的节点之间的边对应的初始权重参数;根据多个节点之间的边对应的初始权重参数,对图网络中多个节点之间的边的权重矩阵进行初始化。

[0155] 其中,样本字段所属文本框的样本位置信息,具体可以为样本字段所属文本框的四个角点各自的坐标。

[0156] 中心坐标,为样本字段所属文本框的中心点的坐标。

[0157] 此处的样本图像,可以为训练数据包括的多个样本图像中,任一包含多个样本字段的样本图像。即本公开实施例中,可以依据任一张包含多个样本字段的样本图像中,多个样本字段所属文本框的样本位置信息,对图网络中多个节点之间的边的权重矩阵进行初始化。

[0158] 可以理解的是,针对一张样本图像中每个样本字段,可以根据样本字段所属文本框的样本位置信息,确定该样本字段所属文本框的中心坐标、宽度和高度,根据多个样本字段所属文本框的中心坐标、宽度和高度,确定图网络中与多个样本字段对应的节点之间的边对应的初始权重参数,进而根据多个节点之间的边对应的初始权重参数,对图网络中多个节点之间的边的权重矩阵进行初始化。

[0159] 在本公开的实施例中,可以采用以下方式,实现根据多个样本字段所属文本框的中心坐标、宽度和高度,确定图网络中与多个样本字段对应的节点之间的边对应的初始权重参数:

[0160] 在多个节点中第一节点指向第二节点的情况下,根据第一节点和第二节点分别对应的样本字段所属文本框的中心坐标,确定第二节点与第一节点对应第一坐标轴的第一坐标差值,以及对应第二坐标轴的第二坐标差值;对第一坐标差值及第二坐标差值作积,得到第一乘积;对第一节点对应的样本字段所属文本框的宽度和高度作积,得到第二乘积;将第一乘积与第二乘积的比值,确定为第一节点与第二节点之间的边的初始权重参数。

[0161] 其中,第一坐标轴和第二坐标轴,可以分别为x轴和y轴。

[0162] 第一坐标差值,为第二节点对应的样本字段所属文本框的中心坐标在第一坐标轴上的坐标值,与第一节点对应的样本字段所属文本框的中心坐标在第一坐标轴上的坐标值之间的差值。

[0163] 第二坐标差值,为第二节点对应的样本字段所属文本框的中心坐标在第二坐标轴上的坐标值,与第一节点对应的样本字段所属文本框的中心坐标在第二坐标轴上的坐标值之间的差值。

[0164] 即,在第一节点指向第二节点的情况下,可以通过以下公式(2),实现根据第一节点对应的样本字段所属文本框的中心坐标、宽度和高度、第二节点对应的样本字段所属文本框的中心坐标,确定图网络中与多个样本字段对应的节点之间的边对应的初始权重参数。

$$[0165] \quad R_{ij} = \frac{|X_j - X_i| * |Y_j - Y_i|}{W_i * H_i} \quad (2)$$

[0166] 其中, R_{ij} 表示第一节点与第二节点之间边的初始权重参数, (X_i, Y_i) 表示第一节点对应的样本字段所属文本框的中心坐标, W_i 为第一节点对应的样本字段所属文本框的宽度, H_i 为第一节点对应的样本字段所属文本框的高度, (X_j, Y_j) 表示第二节点对应的样本字段所属文本框的中心坐标。

[0167] 通过上述方式,实现了在第一节点指向第二节点的情况下,基于第一节点对应的样本字段所属文本框的中心坐标、宽度和高度、第二节点对应的样本字段所属文本框的中

心坐标,计算得到第一节点与第二节点之间边的初始权重参数。通过这种方式,可以计算得到图网络中各节点之间的边的初始权重参数,进而使得能够根据多个节点之间的边对应的初始权重参数,对图网络中多个节点之间的边的权重矩阵进行初始化,提高后续训练得到的目标结构化信息抽取模型的模型性能。

[0168] 下面结合图7,对本公开提供的结构化信息抽取装置进行说明。

[0169] 图7是根据本公开第六实施例的结构化信息抽取装置的结构示意图。

[0170] 如图7所示,本公开提供的结构化信息抽取装置700,包括:第一获取模块701、融合模块702、第一确定模块703以及组织模块704。

[0171] 其中,第一获取模块701,用于获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息;

[0172] 融合模块702,用于将至少一个字段对应的图像特征和文字特征进行融合,以得到至少一个字段对应的融合特征;

[0173] 第一确定模块703,用于基于至少一个字段对应的融合特征,确定至少一个字段对应的目标类别;

[0174] 组织模块704,用于根据至少一个字段对应的目标类别之间的对应关系,对至少一个字段对应的文字信息进行组织,以得到结构化信息。

[0175] 需要说明的是,本实施例提供的结构化信息抽取装置700,可以执行前述实施例的结构化信息抽取方法。其中,结构化信息抽取装置700可以由软件和/或硬件实现,该结构化信息抽取装置700可以配置在电子设备,该电子设备可以包括但不限于终端设备、服务器等,该实施例对电子设备不作具体限定。

[0176] 需要说明的是,前述对于结构化信息抽取方法的实施例的说明,也适用于本公开提供的结构化信息抽取装置,此处不再赘述。

[0177] 本公开实施例提供的结构化信息抽取装置,通过获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息,将至少一个字段对应的图像特征和文字特征进行融合,以得到至少一个字段对应的融合特征,基于至少一个字段对应的融合特征,确定至少一个字段对应的目标类别,根据至少一个字段对应的目标类别之间的对应关系,对至少一个字段对应的文字信息进行组织,以得到结构化信息,由此,只需确定待处理图像中至少一个字段对应的目标类别,即可根据至少一个字段对应的目标类别之间的对应关系,实现对结构化信息的抽取,从而实现了将结构化信息抽取问题简化为分类问题,简化了结构化信息抽取的过程,并且,由于结合待处理图像中至少一个字段的文字特征和图像特征两种特征进行结构化信息抽取,提高了结构化信息抽取的准确性。

[0178] 下面结合图8,对本公开提供的结构化信息抽取装置进行进一步说明。

[0179] 图8是根据本公开第七实施例的结构化信息抽取装置的结构示意图。

[0180] 如图8所示,结构化信息抽取装置800,具体可以包括:第一获取模块801、融合模块802、第一确定模块803以及组织模块804。其中,图8中第一获取模块801、融合模块802、第一确定模块803以及组织模块804与图7中第一获取模块701、融合模块702、第一确定模块703以及组织模块704具有相同功能和结构。

[0181] 在本公开的实施例中,至少一个字段对应的目标类别,为多个属性项类别中的至少一个或多个属性值类别中的至少一个;属性项类别和属性值类别之间具有一一对应关

系；

[0182] 其中,组织模块804,包括:

[0183] 第一查询单元8041,用于根据目标类别为第一属性项类别的第一字段,查询目标类别为具有对应关系的第一属性值类别的第二字段;

[0184] 第一获取单元8042,用于获取第一字段对应的第一文字信息,以及第二字段对应的第二文字信息;

[0185] 第一生成单元8043,用于根据第一文字信息和第二文字信息,生成一条结构化信息。

[0186] 在本公开的实施例中,至少一个字段对应的目标类别,为多个表头类别中的至少一个或多个表格单元类别中的至少一个;表头类别和表格单元类别之间具有一一对应关系;

[0187] 其中,组织模块804,包括:

[0188] 第二查询单元,用于根据目标类别为第一表头类别的第三字段,查询目标类别为具有对应关系的第一表格单元类别的第四字段;

[0189] 第二获取单元,用于获取第三字段对应的第三文字信息,以及第四字段对应的第四文字信息;

[0190] 第二生成单元,用于根据第三文字信息和第四文字信息,生成一条结构化信息。

[0191] 在本公开的实施例中,第二字段的数量为多个,第一生成单元8043,包括:

[0192] 合并子单元,用于对多个第二字段对应的第二文字信息进行合并,得到合并信息;

[0193] 生成子单元,用于根据第一文字信息和合并信息,生成一条结构化信息。

[0194] 在本公开的实施例中,第一确定模块803,包括:

[0195] 第三获取单元8031,用于将至少一个字段对应的融合特征,作为图网络中至少一个节点的节点特征,输入图网络中,以获取图网络输出的至少一个字段对应多个预设类别的概率;其中,至少一个字段与图网络中的节点相对应;

[0196] 第一确定单元8032,用于根据概率,从多个预设类别中,确定至少一个字段对应的目标类别。

[0197] 在本公开的实施例中,图网络包含于结构化信息抽取模型,结构化信息抽取模型,还包括特征提取网络,第一获取模块801,包括:

[0198] 第四获取单元8011,用于获取对待处理图像进行光学字符识别得到的待处理图像中至少一个字段对应的文字信息,以及至少一个字段所属文本框的位置信息;

[0199] 第一特征提取单元8012,用于将待处理图像、待处理图像中至少一个字段对应的文字信息,以及至少一个字段所属文本框的位置信息,输入特征提取网络,以采用特征提取网络进行特征提取,获取待处理图像中至少一个字段对应的图像特征和文字特征。

[0200] 在本公开的实施例中,特征提取网络包括卷积神经网络和循环神经网络;

[0201] 其中,第一特征提取单元8012,包括:

[0202] 第一提取子单元,用于将待处理图像输入卷积神经网络,以采用卷积神经网络对待处理图像进行图像特征提取,得到特征图;

[0203] 第一截取子单元,用于根据至少一个字段所属文本框的位置信息,截取特征图中对应的区域,得到至少一个字段对应的图像特征;

[0204] 第二提取子单元,用于将至少一个字段对应的文字信息输入循环神经网络,以采用循环神经网络对文字信息进行文字特征提取,得到至少一个字段对应的文字特征。

[0205] 在本公开的实施例中,结构化信息抽取模型,还包括池化层,第一特征提取单元8012,还包括:

[0206] 池化处理子单元,用于将图像特征输入池化层,以采用池化层对图像特征进行池化处理,以使至少一个字段对应的图像特征的维度相同。

[0207] 在本公开的实施例中,结构化信息抽取模型,还包括特征融合层,融合模块802,包括:

[0208] 第一处理单元8021,用于将至少一个字段对应的图像特征和文字特征输入特征融合层,以采用特征融合层,对至少一个字段对应的图像特征和文字特征进行拼接,或者对至少一个字段对应的图像特征和文字特征进行克罗内克积运算,得到至少一个字段对应的融合特征。

[0209] 需要说明的是,前述对于结构化信息抽取方法的实施例的说明,也适用于本公开提供的结构化信息抽取装置,此处不再赘述。

[0210] 本公开实施例提供的结构化信息抽取装置,通过获取待处理图像中至少一个字段对应的图像特征、文字特征以及文字信息,将至少一个字段对应的图像特征和文字特征进行融合,以得到至少一个字段对应的融合特征,基于至少一个字段对应的融合特征,确定至少一个字段对应的目标类别,根据至少一个字段对应的目标类别之间的对应关系,对至少一个字段对应的文字信息进行组织,以得到结构化信息,由此,只需确定待处理图像中至少一个字段对应的目标类别,即可根据至少一个字段对应的目标类别之间的对应关系,实现对结构化信息的抽取,从而实现了将结构化信息抽取问题简化为分类问题,简化了结构化信息抽取的过程,并且,由于结合待处理图像中至少一个字段的文字特征和图像特征两种特征进行结构化信息抽取,提高了结构化信息抽取的准确性。

[0211] 根据本公开的实施例,还提供一种用于结构化信息抽取的模型训练装置。

[0212] 下面结合图9,对本公开提供的用于结构化信息抽取的模型训练装置进行说明。

[0213] 图9是根据本公开第八实施例的用于结构化信息抽取的模型训练装置的结构示意图。

[0214] 如图9所示,本公开提供的用于结构化信息抽取的模型训练装置900,包括:第二获取模块901、第二确定模块902以及第三获取模块903。

[0215] 其中,第二获取模块901,用于获取训练数据及初始结构化信息抽取模型;其中,训练数据包括多个样本图像、多个样本图像中至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,至少一个样本字段以对应的样本类别进行标注;

[0216] 第二确定模块902,用于将至少一个样本图像、样本图像中至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,输入初始结构化信息抽取模型,以采用初始结构化信息抽取模型确定至少一个样本字段对应的预测类别;

[0217] 第三获取模块903,用于获取至少一个样本字段对应的预测类别与样本类别之间的差异,以根据差异对初始结构化信息抽取模型的模型参数进行调整,得到用于结构化信息抽取的目标结构化信息抽取模型。

[0218] 其中,目标结构化信息抽取模型用于执行以下步骤:

[0219] 获取待处理图像中至少一个字段对应的图像特征和文字特征;

[0220] 将至少一个字段对应的图像特征和文字特征进行融合,以得到至少一个字段对应的融合特征;

[0221] 基于至少一个字段对应的融合特征,确定至少一个字段对应的目标类别,其中,至少一个字段对应的目标类别之间的对应关系,用于对至少一个字段对应的文字信息进行组织,以得到结构化信息。

[0222] 需要说明的是,本实施例提供的用于结构化信息抽取的模型训练装置900,简称模型训练装置,可以执行前述实施例的用于结构化信息抽取的模型训练方法。其中,模型训练装置可以由软件和/或硬件实现,该模型训练装置可以配置在电子设备,该电子设备可以包括但不限于终端设备、服务器等,该实施例对电子设备不作具体限定。

[0223] 需要说明的是,前述对于用于结构化信息抽取的模型训练方法的实施例的说明,也适用于本公开提供的用于结构化信息抽取的模型训练装置,此处不再赘述。

[0224] 本公开实施例提供的用于结构化信息抽取的模型训练装置,通过获取训练数据及初始结构化信息抽取模型,将至少一个样本图像、样本图像中至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,输入初始结构化信息抽取模型,以采用初始结构化信息抽取模型确定至少一个样本字段对应的预测类别,获取至少一个样本字段对应的预测类别与样本类别之间的差异,以根据差异对初始结构化信息抽取模型的模型参数进行调整,得到用于结构化信息抽取的目标结构化信息抽取模型,实现了基于训练数据,对初始结构化信息抽取模型进行模型训练,得到用于结构化信息抽取的目标结构化信息抽取模型,利用训练后得到的目标结构化信息抽取模型进行结构化信息抽取,能够实现将结构化信息抽取问题简化为分类问题,简化结构化信息抽取的过程,并且,通过结合待处理图像中至少一个字段的文字特征和图像特征两种特征进行结构化信息抽取,提高结构化信息抽取的准确性。

[0225] 下面结合图10,对本公开提供的用于结构化信息抽取的模型训练装置进行进一步说明。

[0226] 图10是根据本公开第九实施例的用于结构化信息抽取的模型训练装置的结构示意图。

[0227] 如图10所示,用于结构化信息抽取的模型训练装置1000,具体可以包括:第二获取模块1001、第二确定模块1002以及第三获取模块1003。其中,图10中第二获取模块1001、第二确定模块1002以及第三获取模块1003与图9中第二获取模块901、第二确定模块902以及第三获取模块903具有相同功能和结构。

[0228] 在本公开的实施例中,至少一个样本字段对应的预测类别、样本类别或目标类别,为多个属性项类别中的至少一个或多个属性值类别中的至少一个,或多个表头类别中的至少一个或多个表格单元类别中的至少一个;属性项类别和属性值类别之间具有一一对应关系;表头类别和表格单元类别之间具有一一对应关系。

[0229] 在本公开的实施例中,初始结构化信息抽取模型包括特征提取网络、特征融合层和图网络;

[0230] 第二确定模块1002,包括:

[0231] 第二特征提取单元,用于将样本图像、样本图像中至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,输入特征提取网络,以采用特征提取网络进行特征提取,获取样本图像中至少一个样本字段对应的样本图像特征和样本文字特征;

[0232] 第二融合单元,用于将至少一个样本字段对应的样本图像特征和样本文字特征输入特征融合层,以采用特征融合层,对至少一个样本字段对应的样本图像特征和样本文字特征进行融合,得到至少一个样本字段对应的样本融合特征;

[0233] 特征编码单元,用于将至少一个样本字段对应的样本融合特征,作为图网络中至少一个节点的节点特征,输入图网络中,以采用图网络对至少一个节点的节点特征进行特征编码,得到至少一个样本字段对应的预测类别。

[0234] 在本公开的实施例中,用于结构化信息抽取的模型训练装置,还包括:

[0235] 第四获取模块1004,用于根据样本图像中多个样本字段所属文本框的样本位置信息,获取多个样本字段所属文本框的中心坐标、宽度和高度;

[0236] 第三确定模块1005,用于根据多个样本字段所属文本框的中心坐标、宽度和高度,确定图网络中与多个样本字段对应的节点之间的边对应的初始权重参数;

[0237] 初始化模块1006,用于根据多个节点之间的边对应的初始权重参数,对图网络中多个节点之间的边的权重矩阵进行初始化。

[0238] 在本公开的实施例中,第三确定模块1005,包括:

[0239] 第二确定单元,用于在多个节点中第一节点指向第二节点的情况下,根据第一节点和第二节点分别对应的样本字段所属文本框的中心坐标,确定第二节点与第一节点对应第一坐标轴的第一坐标差值,以及对应第二坐标轴的第二坐标差值;

[0240] 第二处理单元,用于对第一坐标差值及第二坐标差值作积,得到第一乘积;

[0241] 第三处理单元,用于对第一节点对应的样本字段所属文本框的宽度和高度作积,得到第二乘积;

[0242] 第三确定单元,用于将第一乘积与第二乘积的比值,确定为第一节点与第二节点之间的边的初始权重参数。

[0243] 在本公开的实施例中,第二融合单元,包括:

[0244] 融合子单元,用于采用特征融合层,对至少一个样本字段对应的样本图像特征和样本文字特征进行拼接,或者对至少一个样本字段对应的图像特征和文字特征进行克罗内克积运算,得到至少一个样本字段对应的样本融合特征。

[0245] 在本公开的实施例中,特征提取网络,包括卷积神经网络和训练神经网络;

[0246] 其中,第二特征提取单元,包括:

[0247] 第三提取子单元,用于将样本图像输入卷积神经网络,以采用卷积神经网络对样本图像进行图像特征提取,得到样本特征图;

[0248] 第二截取子单元,用于根据至少一个样本字段所属文本框的样本位置信息,截取样本特征图中对应的区域,得到至少一个样本字段对应的样本图像特征;

[0249] 第四提取子单元,用于将样本图像中至少一个样本字段对应的样本文字信息输入循环神经网络,以采用循环神经网络对样本文字信息进行文字特征提取,得到至少一个样本字段对应的样本文字特征。

[0250] 本公开实施例提供的用于结构化信息抽取的模型训练装置,通过获取训练数据及初始结构化信息抽取模型,将至少一个样本图像、样本图像中至少一个样本字段对应的样本文字信息,以及至少一个样本字段所属文本框的样本位置信息,输入初始结构化信息抽取模型,以采用初始结构化信息抽取模型确定至少一个样本字段对应的预测类别,获取至少一个样本字段对应的预测类别与样本类别之间的差异,以根据差异对初始结构化信息抽取模型的模型参数进行调整,得到用于结构化信息抽取的目标结构化信息抽取模型,实现了基于训练数据,对初始结构化信息抽取模型进行模型训练,得到用于结构化信息抽取的目标结构化信息抽取模型,利用训练后得到的目标结构化信息抽取模型进行结构化信息抽取,能够实现将结构化信息抽取问题简化为分类问题,简化结构化信息抽取的过程,并且,通过结合待处理图像中至少一个字段的文字特征和图像特征两种特征进行结构化信息抽取,提高结构化信息抽取的准确性。

[0251] 基于上述实施例,本公开还提供一种电子设备,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本公开的结构化信息抽取方法,或者执行本公开的用于结构化信息抽取的模型训练方法。

[0252] 基于上述实施例,本公开还提供一种存储有计算机指令的非瞬时计算机可读存储介质,所述计算机指令用于使所述计算机执行本公开实施例公开的结构化信息抽取方法,或者执行本公开实施例公开的用于结构化信息抽取的模型训练方法。

[0253] 基于上述实施例,本公开还提供一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时实现本公开的结构化信息抽取方法的步骤,或者实现本公开的用于结构化信息抽取的模型训练方法的步骤。

[0254] 根据本公开的实施例,本公开还提供了一种电子设备和一种可读存储介质和一种计算机程序产品。

[0255] 图11示出了可以用来实施本公开的实施例的示例电子设备1100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,个人数字处理、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。

[0256] 如图11所示,该电子设备1100可以包括计算单元1101,其可以根据存储在只读存储器(ROM) 1102中的计算机程序或者从存储单元1108加载到随机访问存储器(RAM) 1103中的计算机程序,来执行各种适当的动作和处理。在RAM 1103中,还可存储设备1100操作所需的各种程序和数据。计算单元1101、ROM 1102以及RAM 1103通过总线1104彼此相连。输入/输出(I/O)接口1105也连接至总线1104。

[0257] 设备1100中的多个部件连接至I/O接口1105,包括:输入单元1106,例如键盘、鼠标等;输出单元1107,例如各种类型的显示器、扬声器等;存储单元1108,例如磁盘、光盘等;以及通信单元1109,例如网卡、调制解调器、无线通信收发机等。通信单元1109允许设备1100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。

[0258] 计算单元1101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1101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1101执行上文所描述的各个方法和处理,例如结构化信息抽取方法或用于结构化信息抽取的模型训练方法。例如,在一些实施例中,结构化信息抽取方法或用于结构化信息抽取的模型训练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1108。在一些实施例中,计算机程序的部分或者全部可以经由ROM 1102和/或通信单元1109而被载入和/或安装到设备1100上。当计算机程序加载到RAM 1103并由计算单元1101执行时,可以执行上文描述的结构化信息抽取方法或用于结构化信息抽取的模型训练方法的一个或多个步骤。备选地,在其他实施例中,计算单元1101可以通过其他任何适当的方式(例如,借助于固件)而被配置为执行结构化信息抽取方法或用于结构化信息抽取的模型训练方法。

[0259] 本文中以上描述的系统和技术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。

[0260] 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。

[0261] 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。

[0262] 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。

[0263] 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的实施方式交互)、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)、互联网和区块链网络。

[0264] 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。服务器可以是云服务器,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。

[0265] 其中,需要说明的是,人工智能是研究使计算机来模拟人的某些思维过程和智能行为(如学习、推理、思考、规划等)的学科,既有硬件层面的技术也有软件层面的技术。人工智能硬件技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理等技术;人工智能软件技术主要包括计算机视觉技术、语音识别技术、自然语言处理技术以及机器学习/深度学习、大数据处理技术、知识图谱技术等几大方向。

[0266] 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。

[0267] 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。

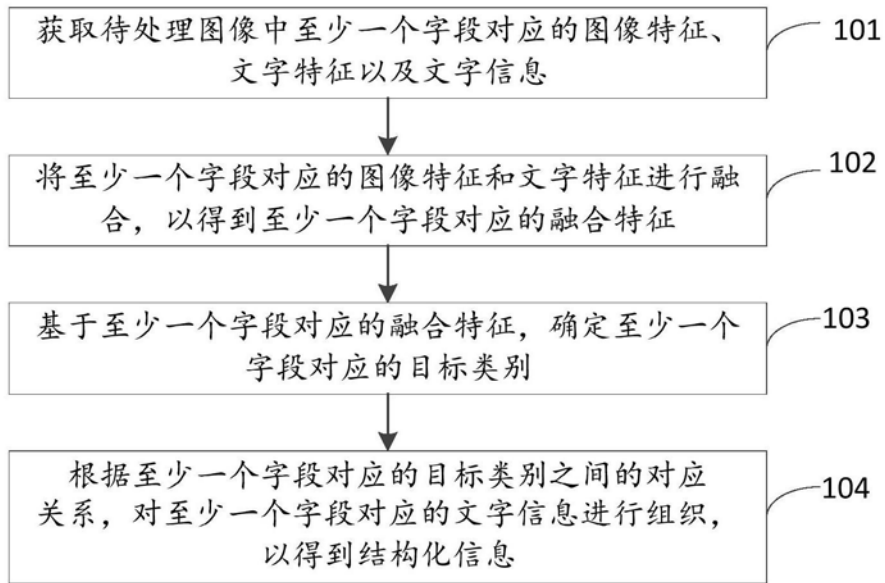


图1

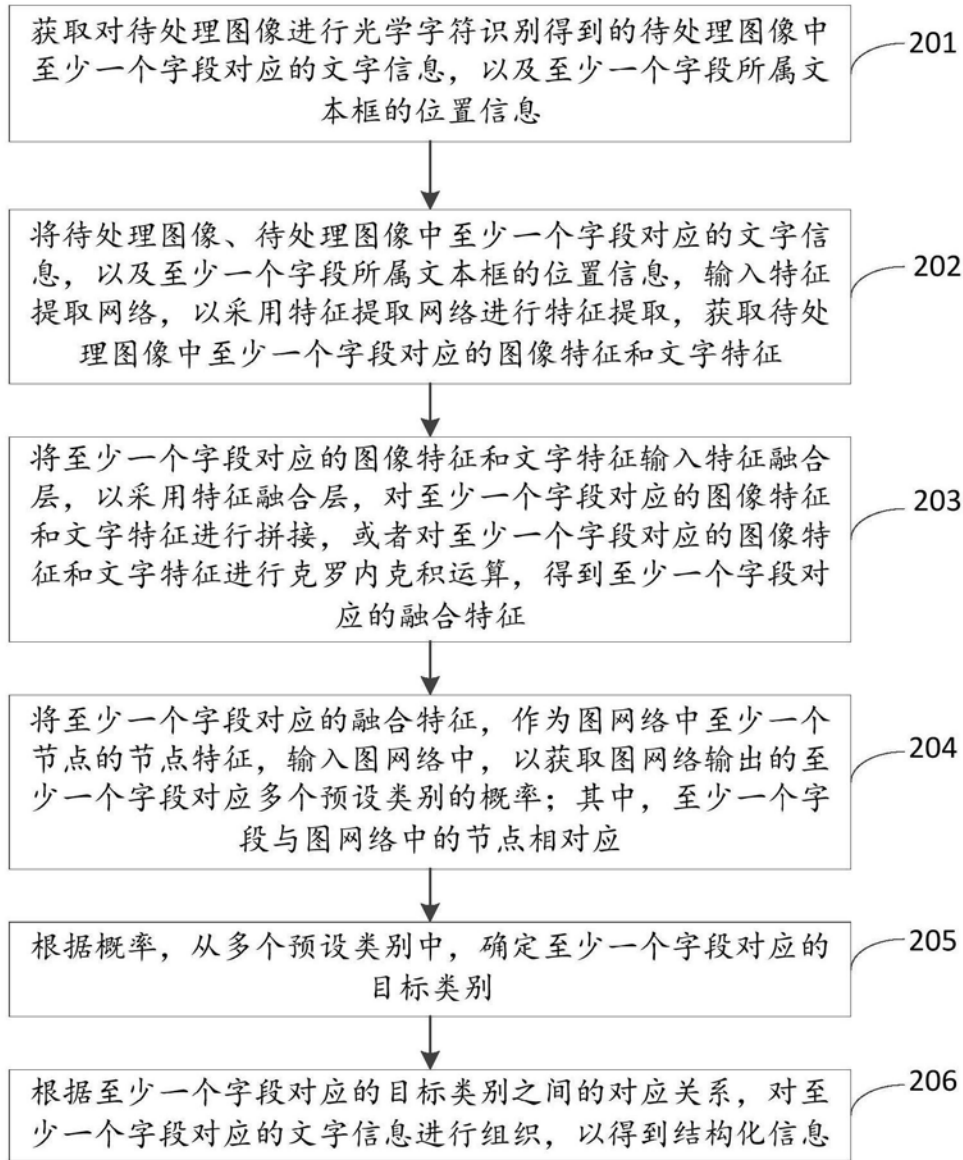


图2

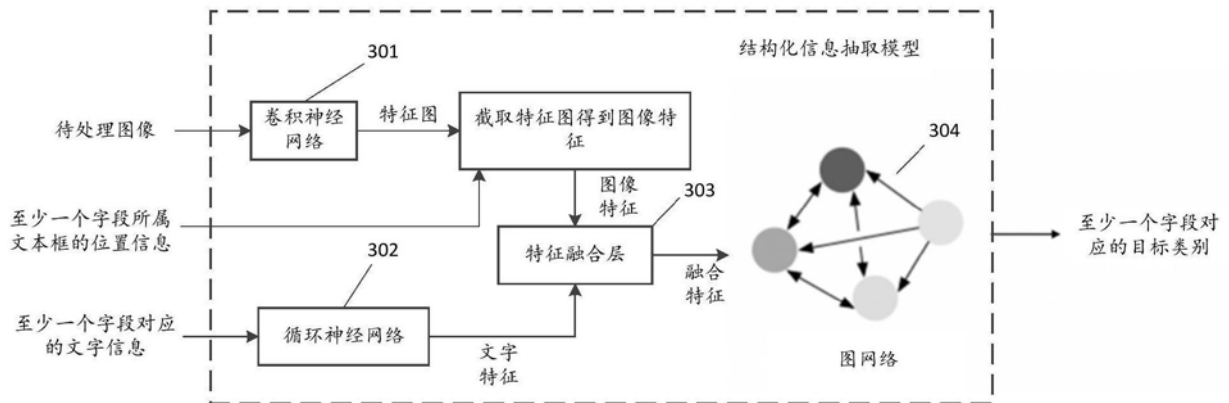


图3

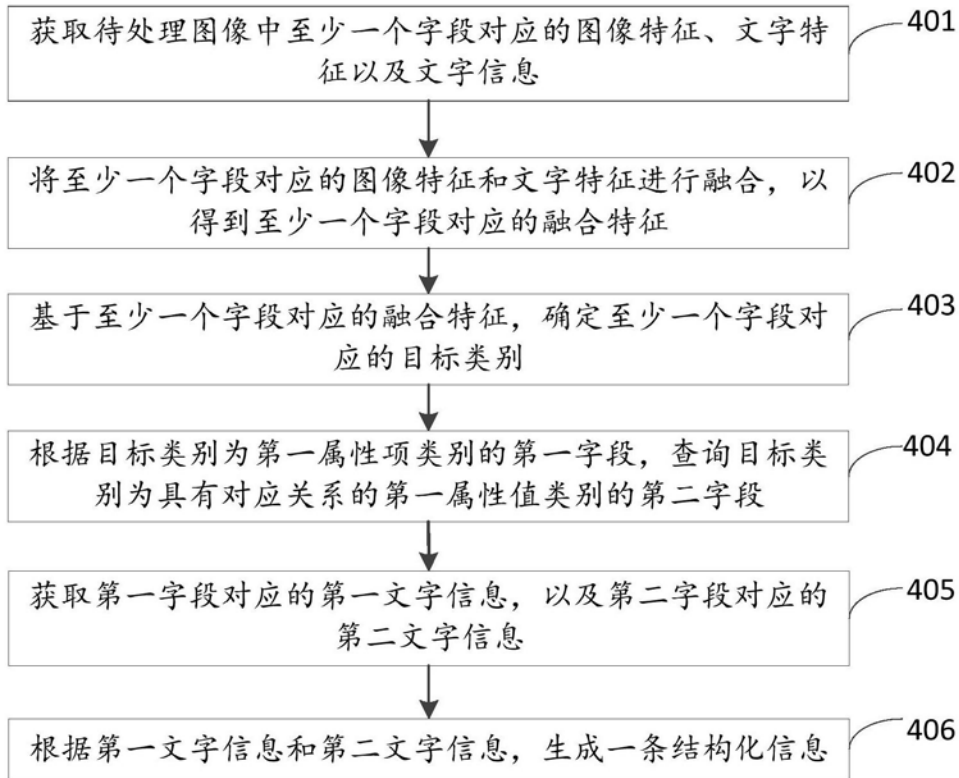


图4

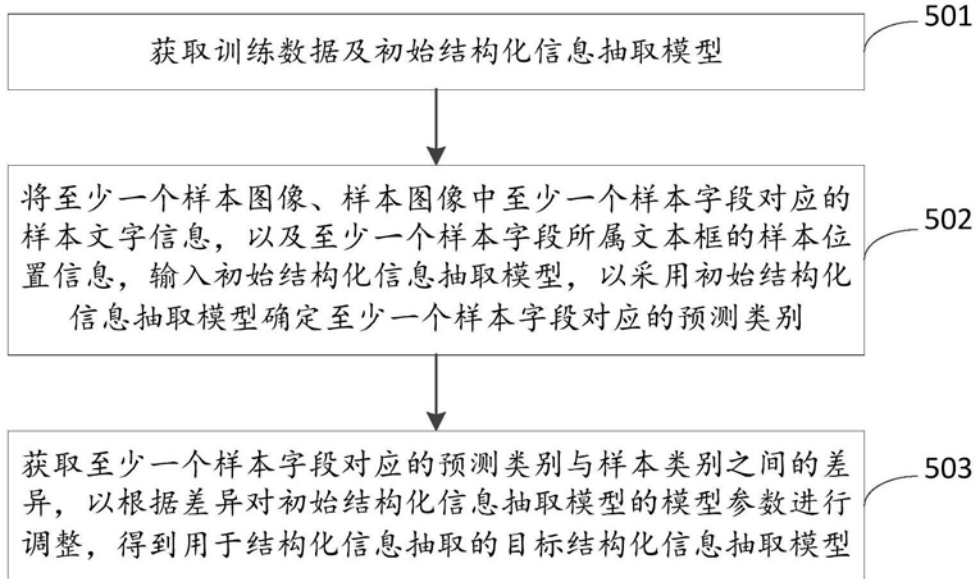


图5

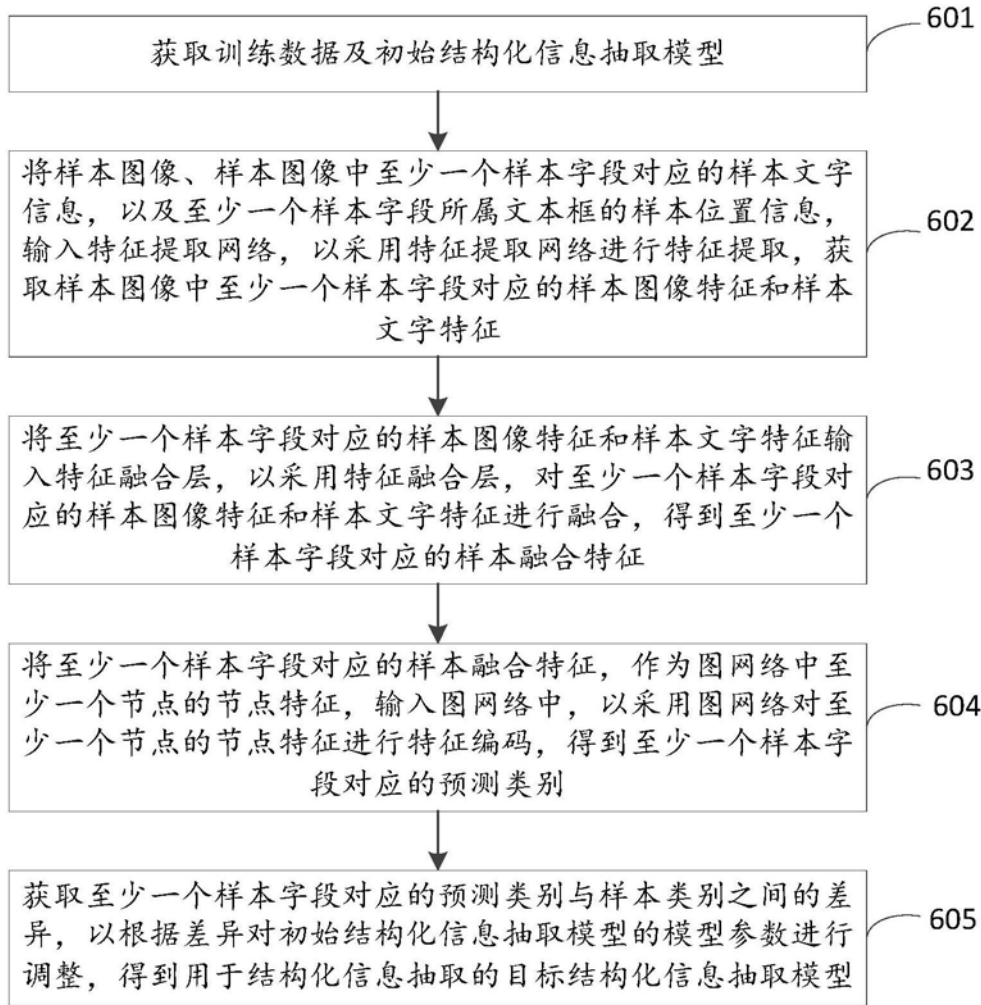


图6

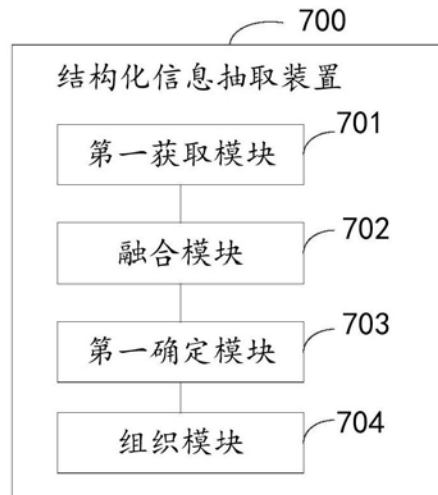


图7

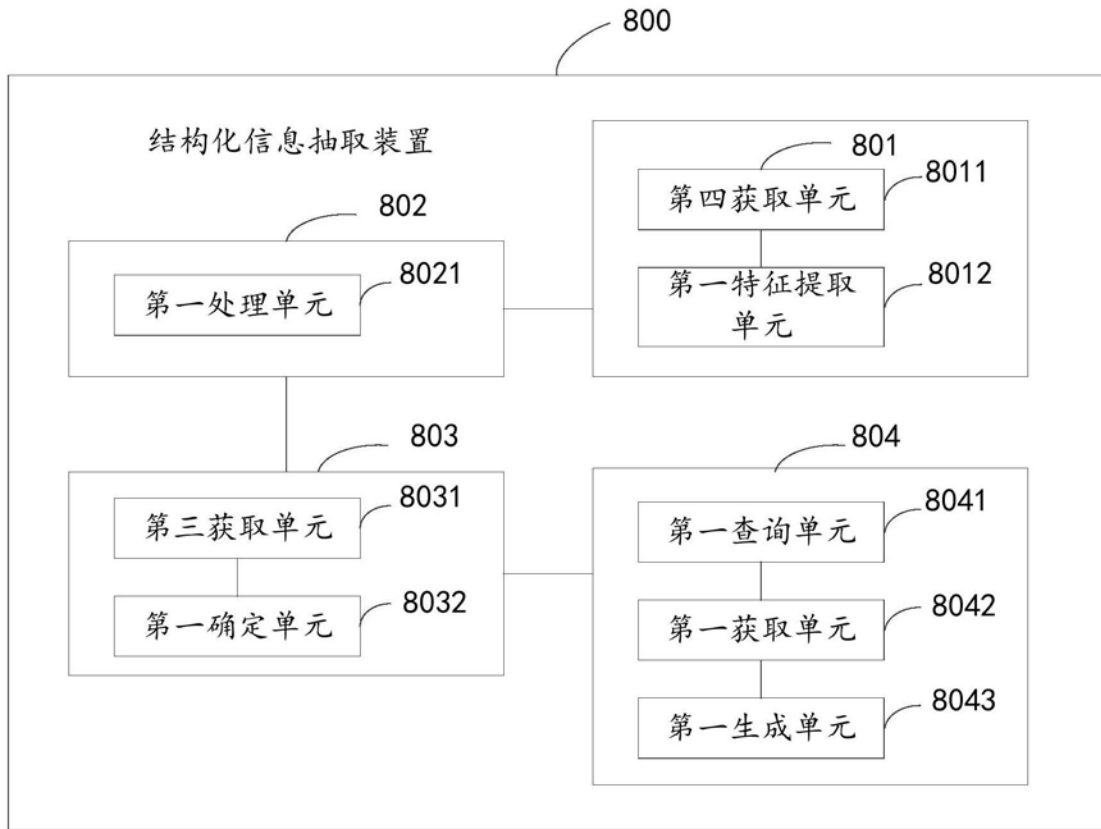


图8



图9

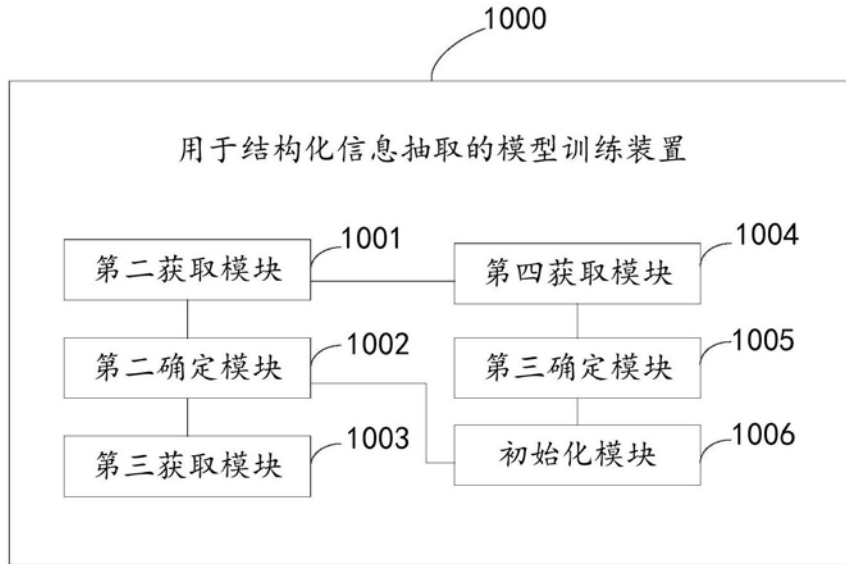


图10

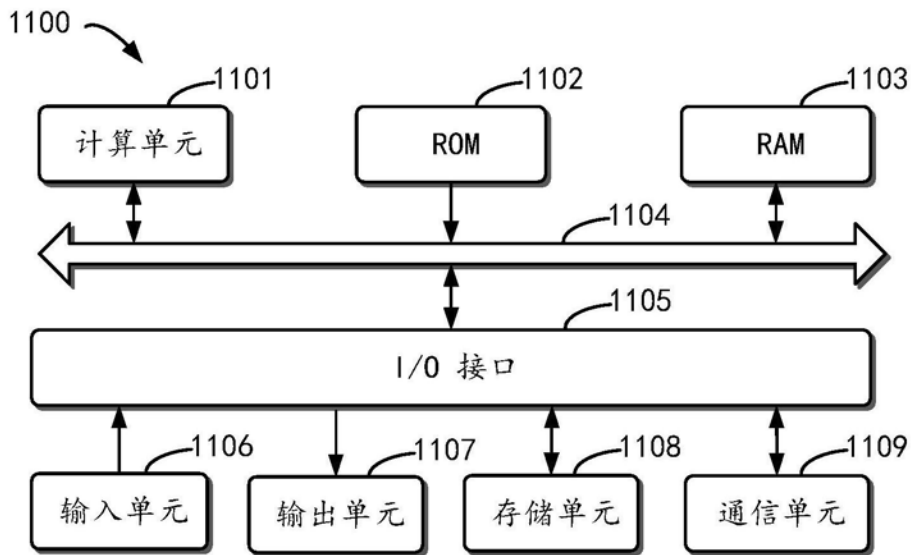


图11