



(12) 发明专利

(10) 授权公告号 CN 114049493 B

(45) 授权公告日 2022.04.01

(21) 申请号 202210024129.0

G06V 10/74 (2022.01)

(22) 申请日 2022.01.11

G06K 9/62 (2022.01)

(65) 同一申请的已公布的文献号

G06F 16/36 (2019.01)

申请公布号 CN 114049493 A

审查员 韦佳黎

(43) 申请公布日 2022.02.15

(73) 专利权人 北京融信数联科技有限公司

地址 100000 北京市海淀区上地信息路26号10层1004室

(72) 发明人 张广志 成立立 于笑博 杨占军 刘畔青

(74) 专利代理机构 北京沃杰永益知识产权代理
事务所(普通合伙) 11905

代理人 杨杰

(51) Int. Cl.

G06V 10/26 (2022.01)

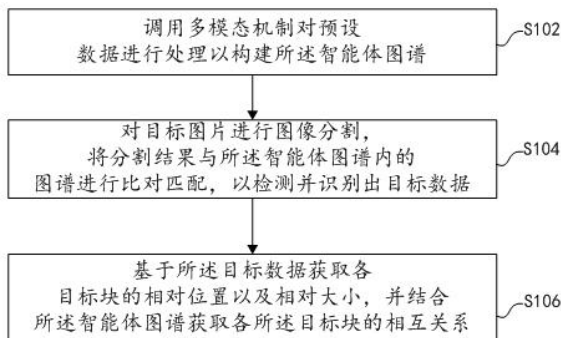
权利要求书3页 说明书9页 附图1页

(54) 发明名称

基于智能体图谱的图像识别方法、系统和可读存储介质

(57) 摘要

本发明公开的一种基于智能体图谱的图像识别方法、系统和可读存储介质,其中方法包括:调用多模态机制对预设数据进行处理以构建所述智能体图谱;对目标图片进行图像分割,将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据;基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系。本发明通过提取多模态数据语义信息,进而通过融合关联,将多模态数据用于智能体图谱构建中,使得智能体图谱内容更加丰富,为认知智能提供多模态本体实体及事理逻辑知识库支撑,依托智能体图谱中的已有认知,进行图像识别,更好地理解图片中的信息。



1. 一种基于智能体图谱的图像识别方法,其特征在于,包括以下步骤:

调用多模态机制对预设数据进行处理以构建所述智能体图谱;

对目标图片进行图像分割,将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据;

基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系;

所述将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据,具体包括:

提取所述分割结果识别每一个所述目标块;

基于每个所述目标块依次与所述智能体图谱内的图片进行比对,并计算相似概率值,其中,若所述相似概率值超过预设第一相似阈值,则判定所述目标块与所述智能体图谱中对应的所述图片为同一物体,否则,判定为不同物体;

将判定为同一物体的所述目标块组合成所述目标数据;

所述基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系,具体包括:

利用区域的矩识别所述目标块于所述目标图片中的方位数据;

计算每个所述目标块的面积并进行大小降序排列,结合所述智能体图谱确定各所述目标块间的所述相互关系,以生成对应的文本化描述;

所述检测并识别出目标数据,还包括:基于分类模型和图像搜索融合技术的检测图像。

2. 根据权利要求1所述的一种基于智能体图谱的图像识别方法,其特征在于,所述调用多模态机制对预设数据进行处理以构建所述智能体图谱,具体包括:

通过多模态知识抽取、多模态知识表示、多模态知识融合以及多模态知识加工来构建所述智能体图谱,其中,

基于预设特征提取模型对所述预设数据进行单模态语义特征提取,而后提取多模态数据的对应关系以完成所述多模态知识抽取;

基于预设嵌入方式将不同类型的所述预设数据投射到同一向量空间中表示,以完成所述多模态知识表示;

基于概念层以及数据层两个层面对所述多模态知识抽取的结果进行链接,以完成所述多模态知识融合;

对所述多模态知识融合后的链接结果调用预设加工方式以完成所述多模态知识加工。

3. 根据权利要求2所述的一种基于智能体图谱的图像识别方法,其特征在于,所述预设嵌入方法包括词嵌入、无监督图嵌入、属性图嵌入以及异构图嵌入;所述加工方式包括本体构建、事件本体构建、知识推理和质量评估。

4. 根据权利要求2所述的一种基于智能体图谱的图像识别方法,其特征在于,所述对目标图片进行图像分割,具体包括:

利用预设算法对所述目标图片进行图像分割得到所述分割结果,其中,所述预设算法为Swin-Transformer算法和/或Self-attention算法;

调用预设参数对所述目标图片进行图像分割得到所述分割结果,其中,所述预设参数包括形状、颜色、亮度以及对比度。

5. 一种基于智能体图谱的图像识别系统,其特征在於,包括存储器和处理器,所述存储器中包括基于智能体图谱的图像识别方法程序,所述基于智能体图谱的图像识别方法程序被所述处理器执行时实现如下步骤:

调用多模态机制对预设数据进行处理以构建所述智能体图谱;

对目标图片进行图像分割,将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据;

基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系;

所述将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据,具体包括:

提取所述分割结果识别每一个所述目标块;

基于每个所述目标块依次与所述智能体图谱内的图片进行比对,并计算相似概率值,其中,若所述相似概率值超过预设第一相似阈值,则判定所述目标块与所述智能体图谱中对应的所述图片为同一物体,否则,判定为不同物体;

将判定为同一物体的所述目标块组合成所述目标数据;

所述基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系,具体包括:

利用区域的矩识别所述目标块于所述目标图片中的方位数据;

计算每个所述目标块的面积并进行大小降序排列,结合所述智能体图谱确定各所述目标块间的所述相互关系,以生成对应的文本化描述;

所述检测并识别出目标数据,还包括:基于分类模型和图像搜索融合技术的检测图像。

6. 根据权利要求5所述的一种基于智能体图谱的图像识别系统,其特征在於,所述调用多模态机制对预设数据进行处理以构建所述智能体图谱,具体包括:

通过多模态知识抽取、多模态知识表示、多模态知识融合以及多模态知识加工来构建所述智能体图谱,其中,

基于预设特征提取模型对所述预设数据进行单模态语义特征提取,而后提取多模态数据的对应关系以完成所述多模态知识抽取;

基于预设嵌入方式将不同类型的所述预设数据投射到同一向量空间中表示,以完成所述多模态知识表示;

基于概念层以及数据层两个层面对所述多模态知识抽取的结果进行链接,以完成所述多模态知识融合;

对所述多模态知识融合后的链接结果调用预设加工方式以完成所述多模态知识加工。

7. 根据权利要求6所述的一种基于智能体图谱的图像识别系统,其特征在於,所述对目标图片进行图像分割,具体包括:

利用预设算法对所述目标图片进行图像分割得到所述分割结果,其中,所述预设算法为Swin-Transformer算法和/或Self-attention算法;

调用预设参数对所述目标图片进行图像分割得到所述分割结果,其中,所述预设参数包括形状、颜色、亮度以及对比度。

8. 一种计算机可读存储介质,其特征在於,所述计算机可读存储介质中包括一种基于

智能体图谱的图像识别方法程序,所述基于智能体图谱的图像识别方法程序被处理器执行时,实现如权利要求1至4中任一项所述的一种基于智能体图谱的图像识别方法的步骤。

基于智能体图谱的图像识别方法、系统和可读存储介质

技术领域

[0001] 本发明涉及图像识别技术领域,更具体的,涉及一种基于智能体图谱的图像识别方法、系统和可读存储介质。

背景技术

[0002] 人工智能已经从计算智能、感知智能,进入到认知智能阶段,认知即个体对知识进行获取、加工、应用的过程,这是人脑的高级信息处理方式;认知智能即让机器具有“读懂语义、逻辑推理、学习判断”的能力,机器认知智能的两个核心是“理解”与“解释”,认知智能的实现需要以知识为驱动力,这涉及到知识表示、语义理解、联想推理、智能问答、情感计算、决策规划等关键技术,随着知识图谱的出现,为认知智能的发展注入了“加速剂”,但目前的知识图谱,主要是以传统本体概念为基础进行知识组织的,而在知识处理领域,这种传统本体概念存在着一些局限性,传统本体对于概念的描述着重对其静态特征的描述,而缺乏对动态特征的描述。

[0003] 为了揭示事件的演化规律和发展模式,旨在传统的知识图谱的基础上,结合事件的演化规律和模式,构建成一个知识与智能库,用于记录和表达客观世界及事件演化规律。

发明内容

[0004] 本发明的目的是提供一种基于智能体图谱的图像识别方法、系统和可读存储介质,可以实现对图片进行多维、立体、有联系的识别,更好地理解图片中不同事物间的关联关系,及其中蕴含的内涵、逻辑。

[0005] 本发明第一方面提供了一种基于智能体图谱的图像识别方法,包括以下步骤:

[0006] 调用多模态机制对预设数据进行处理以构建所述智能体图谱;

[0007] 对目标图片进行图像分割,将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据;

[0008] 基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系。

[0009] 本方案中,所述调用多模态机制对预设数据进行处理以构建所述智能体图谱,具体包括:

[0010] 通过多模态知识抽取、多模态知识表示、多模态知识融合以及多模态知识加工来构建所述智能体图谱,其中,

[0011] 基于预设特征提取模型对所述预设数据进行单模态语义特征提取,而后提取多模态数据的对应关系以完成所述多模态知识抽取;

[0012] 基于预设嵌入方式将不同类型的所述预设数据投射到同一向量空间中表示,以完成所述多模态知识表示;

[0013] 基于概念层以及数据层两个层面对所述多模态知识抽取的结果进行链接,以完成所述多模态知识融合;

[0014] 对所述多模态知识融合后的链接结果调用预设加工方式以完成所述多模态知识加工。

[0015] 本方案中,所述预设嵌入方法包括词嵌入、无监督图嵌入、属性图嵌入以及异构图嵌入;所述加工方式包括本体构建、事件本体构建、知识推理和质量评估。

[0016] 本方案中,所述对目标图片进行图像分割,具体包括:

[0017] 利用预设算法对所述目标图片进行图像分割得到所述分割结果,其中,所述预设算法为Swin-Transformer算法和/或Self-attention算法;

[0018] 调用预设参数对所述目标图片进行图像分割得到所述分割结果,其中,所述预设参数包括形状、颜色、亮度以及对比度。

[0019] 本方案中,所述将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据,具体包括:

[0020] 提取所述分割结果识别每一个所述目标块;

[0021] 基于每个所述目标块依次与所述智能体图谱内的图片进行比对,并计算相似概率值,其中,若所述相似概率值超过预设第一相似阈值,则判定所述目标块与所述智能体图谱中对应的所述图片为同一物体,否则,判定为不同物体;

[0022] 将判定为同一物体的所述目标块组合成所述目标数据。

[0023] 本方案中,所述基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系,具体包括:

[0024] 利用区域的矩识别所述目标块于所述目标图片中的方位数据;

[0025] 计算每个所述目标块的面积并进行大小降序排列,结合所述智能体图谱确定各所述目标块间的所述相互关系,以生成对应的文本化描述。

[0026] 本发明第二方面还提供一种基于智能体图谱的图像识别系统,包括存储器和处理器,所述存储器中包括基于智能体图谱的图像识别方法程序,所述基于智能体图谱的图像识别方法程序被所述处理器执行时实现如下步骤:

[0027] 调用多模态机制对预设数据进行处理以构建所述智能体图谱;

[0028] 对目标图片进行图像分割,将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据;

[0029] 基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系。

[0030] 本方案中,所述调用多模态机制对预设数据进行处理以构建所述智能体图谱,具体包括:

[0031] 通过多模态知识抽取、多模态知识表示、多模态知识融合以及多模态知识加工来构建所述智能体图谱,其中,

[0032] 基于预设特征提取模型对所述预设数据进行单模态语义特征提取,而后提取多模态数据的对应关系以完成所述多模态知识抽取;

[0033] 基于预设嵌入方式将不同类型的所述预设数据投射到同一向量空间中表示,以完成所述多模态知识表示;

[0034] 基于概念层以及数据层两个层面对所述多模态知识抽取的结果进行链接,以完成所述多模态知识融合;

[0035] 对所述多模态知识融合后的链接结果调用预设加工方式以完成所述多模态知识加工。

[0036] 本方案中,所述预设嵌入方法包括词嵌入、无监督图嵌入、属性图嵌入以及异构图嵌入;所述加工方式包括本体构建、事件本体构建、知识推理和质量评估。

[0037] 本方案中,所述对目标图片进行图像分割,具体包括:

[0038] 利用预设算法对所述目标图片进行图像分割得到所述分割结果,其中,所述预设算法为Swin-Transformer算法和/或Self-attention算法;

[0039] 调用预设参数对所述目标图片进行图像分割得到所述分割结果,其中,所述预设参数包括形状、颜色、亮度以及对比度。

[0040] 本方案中,所述将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据,具体包括:

[0041] 提取所述分割结果识别每一个所述目标块;

[0042] 基于每个所述目标块依次与所述智能体图谱内的图片进行比对,并计算相似概率值,其中,若所述相似概率值超过预设第一相似阈值,则判定所述目标块与所述智能体图谱中对应的所述图片为同一物体,否则,判定为不同物体;

[0043] 将判定为同一物体的所述目标块组合成所述目标数据。

[0044] 本方案中,所述基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系,具体包括:

[0045] 利用区域的矩识别所述目标块于所述目标图片中的方位数据;

[0046] 计算每个所述目标块的面积并进行大小降序排列,结合所述智能体图谱确定各所述目标块间的所述相互关系,以生成对应的文本化描述。

[0047] 本发明第三方面提供了一种计算机可读存储介质,所述计算机可读存储介质中包括机器的一种基于智能体图谱的图像识别方法程序,所述基于智能体图谱的图像识别方法程序被处理器执行时,实现如上述任一项所述的一种基于智能体图谱的图像识别方法的步骤。

[0048] 本发明公开的一种基于智能体图谱的图像识别方法、系统和可读存储介质,通过构建智能体图谱,形成对世界的认知,依托智能体图谱的认知,可以实现对图片进行多维、立体、有联系的识别,更好地理解图片中不同事物间的关联关系,及其中蕴含的内涵、逻辑;通过提取多模态数据语义信息,进而通过融合关联,将多模态数据用于智能体图谱构建中,使得智能体图谱内容更加丰富,为认知智能提供多模态本体实体及事理逻辑知识库支撑,依托智能体图谱中的已有认知,进行图像识别,更好地理解图片中的信息。

附图说明

[0049] 图1示出了本发明一种基于智能体图谱的图像识别方法的流程图;

[0050] 图2示出了本发明一种基于智能体图谱的图像识别系统的框图。

具体实施方式

[0051] 为了能够更清楚地理解本发明的上述目的、特征和优点,下面结合附图和具体实施方式对本发明进行进一步的详细描述。需要说明的是,在不冲突的情况下,本申请的实施

例及实施例中的特征可以相互组合。

[0052] 在下面的描述中阐述了很多具体细节以便于充分理解本发明,但是,本发明还可以采用其他不同于在此描述的方式来实施,因此,本发明的保护范围并不受下面公开的具体实施例的限制。

[0053] 图1示出了本申请一种基于智能体图谱的图像识别方法的流程图。

[0054] 如图1所示,本申请公开了一种基于智能体图谱的图像识别方法,包括以下步骤:

[0055] S102,调用多模态机制对预设数据进行处理以构建所述智能体图谱;

[0056] S104,对目标图片进行图像分割,将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据;

[0057] S106,基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系。

[0058] 需要说明的是,首先通过所述多模态知识抽取、所述多模态知识表示、所述多模态知识融合以及所述多模态知识加工,构建所述智能体图谱,而后对所述目标图片进行分割,将分割后的所述分割结果与所述智能体图谱中的图片进行比对匹配,以识别所述目标数据,而后计算所述目标数据中各所述目标块的相对位置以及相对大小,进行排序同时生成对应的所述文本化描述,依托所述智能体图谱中的已有认知,进行图像识别,以更好地理解图片中的信息。

[0059] 根据本发明实施例,所述调用多模态机制对预设数据进行处理以构建所述智能体图谱,具体包括:

[0060] 通过多模态知识抽取、多模态知识表示、多模态知识融合以及多模态知识加工来构建所述智能体图谱,其中,

[0061] 基于预设特征提取模型对所述预设数据进行单模态语义特征提取,而后提取多模态数据的对应关系以完成所述多模态知识抽取;

[0062] 基于预设嵌入方式将不同类型的所述预设数据投射到同一向量空间中表示,以完成所述多模态知识表示;

[0063] 基于概念层以及数据层两个层面对所述多模态知识抽取的结果进行链接,以完成所述多模态知识融合;

[0064] 对所述多模态知识融合后的链接结果调用预设加工方式以完成所述多模态知识加工。

[0065] 需要说明的是,基于所述预设数据的文本、图像、音频以及视频等数据特征构建所述预设特征提取模型,基于所述预设特征提取模型完成所述单模态语义特征提取,其中,包括图像数据语义特征提取、音频数据语义特征提取、视频数据语义特征提取、文本数据语义特征提取等,而后基于提取到的单模态语义特征建立多模态语义特征间的依赖关系,进而提取所述多模态数据的对应关系,以完成所述多模态知识抽取。

[0066] 进一步地,所述预设嵌入方法包括词嵌入、无监督图嵌入、属性图嵌入以及异构图嵌入,故基于所述词嵌入、所述无监督图嵌入、所述属性图嵌入以及所述异构图嵌入等方式,将不同类型数据投射到同一向量空间中进行表示,实现跨模态的所述多模态知识表示。

[0067] 进一步地,从所述概念层和所述数据层两方面,通过对齐、关联、合并等方式,将多个多模态数据源中的本体、实体、事件本体、事件进行链接,形成一个更加统一、稠密的图

谱,所述概念层的知识融合主要表现为本体对齐、事件本体对齐,是指确定本体概念、本体关系、本体属性等本体之间映射关系,以及确定事件本体、事理关系、事件属性的过程;所述数据层的知识融合主要表现为共指消解和实体对齐、具体事件对齐,共指消解意在将同一信息源中同一实体、同一事件的不同标签统一,实现消歧的目的。

[0068] 进一步地,所述加工方式包括本体构建、事件本体构建、知识推理和质量评估,其中,所述本体构建是图谱内实体连通的语义基础,所述事件本体构建是图谱内事件连通的语义基础,所述知识推理是通过已有实体间关系、已有事件间关系的计算,找到实体间新关联、事件间新关联,是图谱更新的重要手段,所述质量评估是知识加工最后的“质检”环节,确保图谱的合理性。

[0069] 根据本发明实施例,所述对目标图片进行图像分割,具体包括:

[0070] 利用预设算法对所述目标图片进行图像分割得到所述分割结果,其中,所述预设算法为Swin-Transformer算法和/或Self-attention算法;

[0071] 调用预设参数对所述目标图片进行图像分割得到所述分割结果,其中,所述预设参数包括形状、颜色、亮度以及对比度。

[0072] 需要说明的是,利用形状、颜色、亮度、对比度,对所述目标图片进行从大到小、从粗到细、多层次的轮廓分割,而后图像分割采用Swin-Transformer算法和/或Self-attention算法进行分割,其中,所述Swin-Transformer算法以及所述Self-attention算法为本领域技术人员根据实际需要选择的现有技术的算法,以满足图像分割的需要,本发明不再一一赘述。

[0073] 根据本发明实施例,所述将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据,具体包括:

[0074] 提取所述分割结果识别每一个所述目标块;

[0075] 基于每个所述目标块依次与所述智能体图谱内的图片进行比对,并计算相似概率值,其中,若所述相似概率值超过预设第一相似阈值,则判定所述目标块与所述智能体图谱中对应的所述图片为同一物体,否则,判定为不同物体;

[0076] 将判定为同一物体的所述目标块组合成所述目标数据。

[0077] 需要说明的是,基于所述目标块与所述智能体图谱内图片进行比对,通过预设相似度计算法则计算所述相似概率值,基于得到的所述相似概率值判断当前所述目标块与所述智能体图谱内的图片是否为同一物体,其中,若所述相似概率值超过预设第一相似阈值,则判定所述目标块与所述智能体图谱中对应的所述图片为同一物体,否则,判定为不同物体,优选地,所述第一相似阈值可取为92%,进一步地,所述相似度计算法则可以应用均值哈希算法,由于均值哈希算法为本领域技术人员根据实际需要选择的现有技术的算法,以满足相似度计算的需要,本发明不再一一赘述。

[0078] 根据本发明实施例,所述基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系,具体包括:

[0079] 利用区域的矩识别所述目标块于所述目标图片中的方位数据;

[0080] 计算每个所述目标块的面积并进行大小降序排列,结合所述智能体图谱确定各所述目标块间的所述相互关系,以生成对应的文本化描述。

[0081] 需要说明的是,所述目标图片中物体方位的确定需要使用所述区域的矩,具体地,

采用中心矩反映区域本身的所述方位数据,通过计算分割出的每个所述目标块的面积,进行面积大小排序,再根据所述智能体图谱中,对各所述目标块间的所述相互关系的定义,确定物体间的联系及主次关系,以生成图像的所述文本化描述。优选地,生成文本化描述采用 Multimodal RNN模型。

[0082] 值得一提的是,所述多模态知识抽取具体包括:

[0083] 基于多模态数据进行实体抽取、实体属性抽取以及实体关系抽取;

[0084] 基于多模态数据进行事件抽取。

[0085] 需要说明的是,所述实体抽取具体指在所述多模态数据源中识别出特定的元素标签,并与预存标签相链接,其中,所述预存标签为所述实体对应的标签;所述实体关系抽取意在找到多模态数据源中实体间的关系,可分为全局抽取和局部抽取;所述实体属性抽取可以看作实体和其属性值之间的关系;所述事件抽取是将所述多模态数据源中的事件信息(事件的时间、地点、人物、原因、结果等)抽取,并结构化地表现出来,包括开放域或限定域事件抽取,事理关系抽取(包含事件顺承、因果、上下位关系抽取等)。

[0086] 值得一提的是,所述检测并识别出目标数据,还包括:基于分类模型和图像搜索融合技术的检测图像。

[0087] 需要说明的是,在对所述目标图片基于所述智能体图谱进行识别时,当识别不出来图片时,可以基于分类模型的物品类别预测,即训练GoogLeNet V1网络预测查询图像类别,使用softmax分类函数,得到每一种可能的类别都有一个所属的概率;基于图像搜索技术的物品类别预测,即使用GoogLeNet V1网络输出的特征在整个图像库中检索最相似的“Top30”图像,对这“30”幅图像所在的类别进行加权排序,同样得到每一种可能的类别都有一个所属的概率;融合方法,将分类模型得到的每类的概率和图像搜索得到的每类的概率加权平均进行降序排列,将位于第一位的图片结果作为该所述目标图片的识别结果。

[0088] 图2示出了本发明一种基于智能体图谱的图像识别系统的框图。

[0089] 如图2所示,本发明公开了一种基于智能体图谱的图像识别系统,包括存储器和处理器,所述存储器中包括基于智能体图谱的图像识别方法程序,所述基于智能体图谱的图像识别方法程序被所述处理器执行时实现如下步骤:

[0090] 调用多模态机制对预设数据进行处理以构建所述智能体图谱;

[0091] 对目标图片进行图像分割,将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据;

[0092] 基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系。

[0093] 需要说明的是,首先通过所述多模态知识抽取、所述多模态知识表示、所述多模态知识融合以及所述多模态知识加工,构建所述智能体图谱,而后对所述目标图片进行分割,将分割后的所述分割结果与所述智能体图谱中的图片进行比对匹配,以识别所述目标数据,而后计算所述目标数据中各所述目标块的相对位置以及相对大小,进行排序同时生成对应的所述文本化描述,依托所述智能体图谱中的已有认知,进行图像识别,以更好地理解图片中的信息。

[0094] 根据本发明实施例,所述调用多模态机制对预设数据进行处理以构建所述智能体图谱,具体包括:

[0095] 通过多模态知识抽取、多模态知识表示、多模态知识融合以及多模态知识加工来构建所述智能体图谱,其中,

[0096] 基于预设特征提取模型对所述预设数据进行单模态语义特征提取,而后提取多模态数据的对应关系以完成所述多模态知识抽取;

[0097] 基于预设嵌入方式将不同类型的所述预设数据投射到同一向量空间中表示,以完成所述多模态知识表示;

[0098] 基于概念层以及数据层两个层面对所述多模态知识抽取的结果进行链接,以完成所述多模态知识融合;

[0099] 对所述多模态知识融合后的链接结果调用预设加工方式以完成所述多模态知识加工。

[0100] 需要说明的是,基于所述预设数据的文本、图像、音频以及视频等数据特征构建所述预设特征提取模型,基于所述预设特征提取模型完成所述单模态语义特征提取,其中,包括图像数据语义特征提取、音频数据语义特征提取、视频数据语义特征提取、文本数据语义特征提取等,而后基于提取到的单模态语义特征建立多模态语义特征间的依赖关系,进而提取所述多模态数据的对应关系,以完成所述多模态知识抽取。

[0101] 进一步地,所述预设嵌入方法包括词嵌入、无监督图嵌入、属性图嵌入以及异构图嵌入,故基于所述词嵌入、所述无监督图嵌入、所述属性图嵌入以及所述异构图嵌入等方式,将不同类型数据投射到同一向量空间中进行表示,实现跨模态的所述多模态知识表示。

[0102] 进一步地,从所述概念层和所述数据层两方面,通过对齐、关联、合并等方式,将多个多模态数据源中的本体、实体、事件本体、事件进行链接,形成一个更加统一、稠密的图谱,所述概念层的知识融合主要表现为本体对齐、事件本体对齐,是指确定本体概念、本体关系、本体属性等本体之间映射关系,以及确定事件本体、事理关系、事件属性的过程;所述数据层的知识融合主要表现为共指消解和实体对齐、具体事件对齐,共指消解意在将同一信息源中同一实体、同一事件的不同标签统一,实现消歧的目的。

[0103] 进一步地,所述加工方式包括本体构建、事件本体构建、知识推理和质量评估,其中,所述本体构建是图谱内实体连通的语义基础,所述事件本体构建是图谱内事件连通的语义基础,所述知识推理是通过已有实体间关系、已有事件间关系的计算,找到实体间新关联、事件间新关联,是图谱更新的重要手段,所述质量评估是知识加工最后的“质检”环节,确保图谱的合理性。

[0104] 根据本发明实施例,所述对目标图片进行图像分割,具体包括:

[0105] 利用预设算法对所述目标图片进行图像分割得到所述分割结果,其中,所述预设算法为Swin-Transformer算法和/或Self-attention算法;

[0106] 调用预设参数对所述目标图片进行图像分割得到所述分割结果,其中,所述预设参数包括形状、颜色、亮度以及对比度。

[0107] 需要说明的是,利用形状、颜色、亮度、对比度,对所述目标图片进行从大到小、从粗到细、多层次的轮廓分割,而后图像分割采用Swin-Transformer算法和/或Self-attention算法进行分割,其中,所述Swin-Transformer算法以及所述Self-attention算法为本领域技术人员根据实际需要选择的现有技术的算法,以满足图像分割的需要,本发明不再一一赘述。

[0108] 根据本发明实施例,所述将分割结果与所述智能体图谱内的图谱进行比对匹配,以检测并识别出目标数据,具体包括:

[0109] 提取所述分割结果识别每一个所述目标块;

[0110] 基于每个所述目标块依次与所述智能体图谱内的图片进行比对,并计算相似概率值,其中,若所述相似概率值超过预设第一相似阈值,则判定所述目标块与所述智能体图谱中对应的所述图片为同一物体,否则,判定为不同物体;

[0111] 将判定为同一物体的所述目标块组合成所述目标数据。

[0112] 需要说明的是,基于所述目标块与所述智能体图谱内图片进行比对,通过预设相似度计算法则计算所述相似概率值,基于得到的所述相似概率值判断当前所述目标块与所述智能体图谱内的图片是否为同一物体,其中,若所述相似概率值超过预设第一相似阈值,则判定所述目标块与所述智能体图谱中对应的所述图片为同一物体,否则,判定为不同物体,优选地,所述第一相似阈值可取为92%,进一步地,所述相似度计算法则可以应用均值哈希算法,由于均值哈希算法为本领域技术人员根据实际需要选择的现有技术的算法,以满足相似度计算的需要,本发明不再一一赘述。

[0113] 根据本发明实施例,所述基于所述目标数据获取各目标块的相对位置以及相对大小,并结合所述智能体图谱获取各所述目标块的相互关系,具体包括:

[0114] 利用区域的矩识别所述目标块于所述目标图片中的方位数据;

[0115] 计算每个所述目标块的面积并进行大小降序排列,结合所述智能体图谱确定各所述目标块间的所述相互关系,以生成对应的文本化描述。

[0116] 需要说明的是,所述目标图片中物体方位的确定需要使用所述区域的矩,具体地,采用中心矩反映区域本身的所述方位数据,通过计算分割出的每个所述目标块的面积,进行面积大小排序,再根据所述智能体图谱中,对各所述目标块间的所述相互关系的定义,确定物体间的联系及主次关系,以生成图像的所述文本化描述。优选地,生成文本化描述采用Multimodal RNN模型。

[0117] 值得一提的是,所述多模态知识抽取具体包括:

[0118] 基于多模态数据进行实体抽取、实体属性抽取以及实体关系抽取;

[0119] 基于多模态数据进行事件抽取。

[0120] 需要说明的是,所述实体抽取具体指在所述多模态数据源中识别出特定的元素标签,并与预存标签相链接,其中,所述预存标签为所述实体对应的标签;所述实体关系抽取意在找到多模态数据源中实体间的关系,可分为全局抽取和局部抽取;所述实体属性抽取可以看作实体和其属性值之间的关系;所述事件抽取是将所述多模态数据源中的事件信息(事件的时间、地点、人物、原因、结果等)抽取,并结构化地表现出来,包括开放域或限定域事件抽取,事理关系抽取(包含事件顺承、因果、上下位关系抽取等)。

[0121] 值得一提的是,所述检测并识别出目标数据,还包括:基于分类模型和图像搜索融合技术的检测图像。

[0122] 需要说明的是,在对所述目标图片基于所述智能体图谱进行识别时,当识别不出来图片时,可以基于分类模型的物品类别预测,即训练GoogLeNet V1网络预测查询图像类别,使用softmax分类函数,得到每一种可能的类别都有一个所属的概率;基于图像搜索技术的物品类别预测,即使用GoogLeNet V1网络输出的特征在整个图像库中检索最相似的

“Top30”图像,对这“30”幅图像所在的类别进行加权排序,同样得到每一种可能的类别都有一个所属的概率;融合方法,将分类模型得到的每类的概率和图像搜索得到的每类的概率加权平均进行降序排列,将位于第一位的图片结果作为该所述目标图片的识别结果。

[0123] 本发明第三方面提供了一种计算机可读存储介质,所述计算机可读存储介质中包括一种基于智能体图谱的图像识别方法程序,所述基于智能体图谱的图像识别方法程序被处理器执行时,实现如上述任一项所述的一种基于智能体图谱的图像识别方法的步骤。

[0124] 本发明公开的一种基于智能体图谱的图像识别方法、系统和可读存储介质,通过构建智能体图谱,形成对世界的认知,依托智能体图谱的认知,可以实现对图片进行多维、立体、有联系的识别,更好地理解图片中不同事物间的关联关系,及其中蕴含的内涵、逻辑;通过提取多模态数据语义信息,进而通过融合关联,将多模态数据用于智能体图谱构建中,使得智能体图谱内容更加丰富,为认知智能提供多模态本体实体及事理逻辑知识库支撑,依托智能体图谱中的已有认知,进行图像识别,更好地理解图片中的信息。

[0125] 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。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,如:多个单元或组件可以结合,或可以集成到另一个系统,或一些特征可以忽略,或不执行。另外,所显示或讨论的各组成部分相互之间的耦合、或直接耦合、或通信连接可以是通过一些接口,设备或单元的间接耦合或通信连接,可以是电性的、机械的或其它形式的。

[0126] 上述作为分离部件说明的单元可以是、或也可以不是物理上分开的,作为单元显示的部件可以是、或也可以不是物理单元;既可以位于一个地方,也可以分布到多个网络单元上;可以根据实际的需要选择其中的部分或全部单元来实现本实施例方案的目的。

[0127] 另外,在本发明各实施例中的各功能单元可以全部集成在一个处理单元中,也可以是各单元分别单独作为一个单元,也可以两个或两个以上单元集成在一个单元中;上述集成的单元既可以采用硬件的形式实现,也可以采用硬件加软件功能单元的形式实现。

[0128] 本领域普通技术人员可以理解:实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成,前述的程序可以存储于计算机可读取存储介质中,该程序在执行时,执行包括上述方法实施例的步骤;而前述的存储介质包括:移动存储设备、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0129] 或者,本发明上述集成的单元如果以软件功能模块的形式实现并作为独立的产品销售或使用,也可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明实施例的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机、服务器、或者网络设备)执行本发明各个实施例所述方法的全部或部分。而前述的存储介质包括:移动存储设备、ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

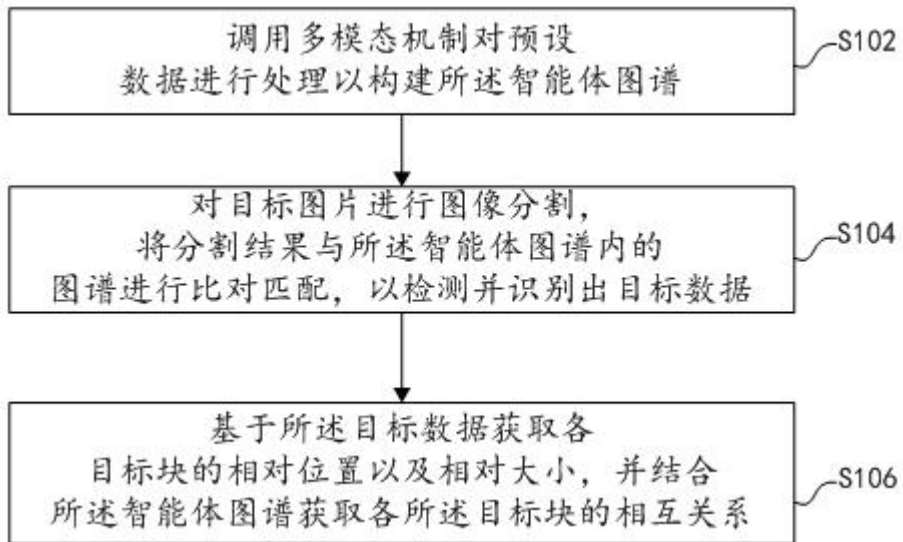


图1

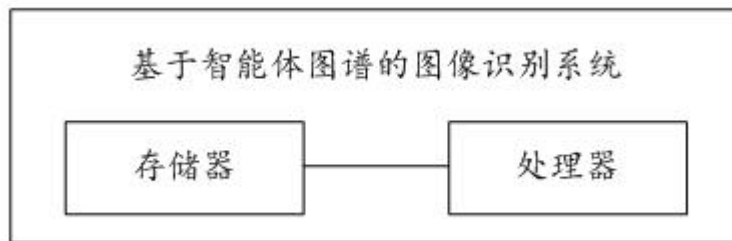


图2