



(12) 发明专利

(10) 授权公告号 CN 111950528 B

(45) 授权公告日 2023. 10. 31

(21) 申请号 202010907711.2

G06V 10/82 (2022.01)

(22) 申请日 2020.09.02

G06N 3/0464 (2023.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 111950528 A

(56) 对比文件

CN 111428875 A, 2020.07.17

US 2013322733 A1, 2013.12.05

(43) 申请公布日 2020.11.17

赵琪; 孙立双; 袁阳. 基于神经网络的无人机拍摄图像识别. 中国科技论文. 2019, (11), 全文.

(73) 专利权人 北京猿力未来科技有限公司

地址 100102 北京市朝阳区广顺南大街8号院1号楼6层F01-03、05-10单元

审查员 刘林

(72) 发明人 康少杰 郭常圳

(74) 专利代理机构 北京智信禾专利代理有限公司

11637

专利代理师 刘晓楠

(51) Int. Cl.

G06V 30/19 (2022.01)

G06V 30/40 (2022.01)

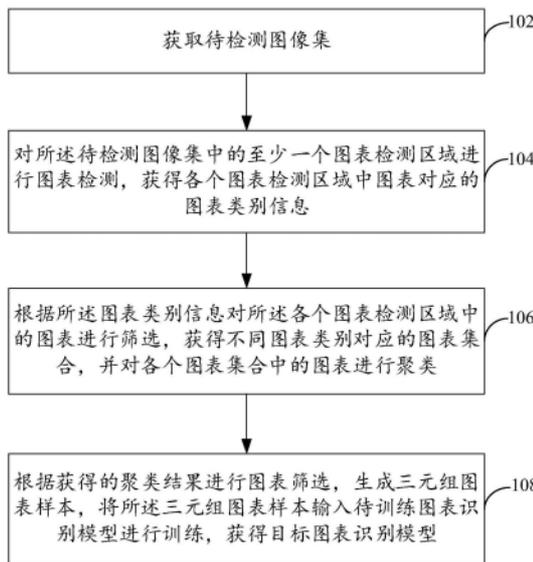
权利要求书4页 说明书17页 附图10页

(54) 发明名称

图表识别模型训练方法以及装置

(57) 摘要

本说明书实施例提供图表识别模型训练方法以及装置,其中所述图表识别模型训练方法包括:获取待检测图像集,对所述待检测图像集中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息,根据所述图表类别信息对所述各个图表检测区域中的图表进行筛选,获得不同图表类别对应的图表集合,并对各个图表集合中的图表进行聚类,根据获得的聚类结果进行图表筛选,生成三元组图表样本,将所述三元组图表样本输入待训练图表识别模型进行训练,获得目标图表识别模型,其中,所述三元组图表样本包括:目标图表、所述目标图表的正样本以及所述目标图表的负样本。



1. 一种图表识别模型训练方法,其特征在于,包括:

获取待检测图像集;

对所述待检测图像集中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息;

根据所述图表类别信息对所述各个图表检测区域中的图表进行筛选,获得不同图表类别对应的图表集合,并对各个图表集合中的图表进行聚类;

根据获得的聚类结果进行图表筛选,生成三元组图表样本,将所述三元组图表样本输入待训练图表识别模型进行训练,获得目标图表识别模型,其中,所述三元组图表样本包括:目标图表、所述目标图表的正样本以及所述目标图表的负样本。

2. 根据权利要求1所述的图表识别模型训练方法,其特征在于,所述对各个图表集合中的图表进行聚类之后,所述根据获得的聚类结果进行图表筛选之前,还包括:

对所述聚类结果中各聚类类别下的图表进行相似性标注;

根据标注结果对所述各聚类类别下的图表进行清洗。

3. 根据权利要求1所述的图表识别模型训练方法,其特征在于,所述对所述待检测图像中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息,包括:

将所述待检测图像输入图表检测模型,确定所述待检测图像中的至少一个图表检测区域;

对所述至少一个图表检测区域进行图表信息识别,获得所述各个图表检测区域中图表对应的图表类别信息。

4. 根据权利要求3所述的图表识别模型训练方法,其特征在于,所述图表检测模型包括特征提取层、候选区域网络层、空间金字塔池化层、池化层;

相应的,所述对所述待检测图像中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息,包括:

将所述待检测图像输入至所述特征提取层,提取所述待检测图像的特征图像;

将所述特征图像输入至所述候选区域网络层,获取所述特征图像对应的至少一个图表检测区域;

将所述特征图像和所述至少一个图表检测区域输入至所述空间金字塔池化层,获得所述至少一个图表检测区域对应的特征向量;

将所述特征向量输入至所述池化层进行池化处理,获得所述至少一个图表检测区域对应的图表类别信息。

5. 根据权利要求1所述的图表识别模型训练方法,其特征在于,所述对各个图表集合中的图表进行聚类,包括:

将所述各个图表集合中的图表输入图表聚类模型,获得所述各个图表集合中图表的聚类结果。

6. 根据权利要求5所述的图表识别模型训练方法,其特征在于,所述图表聚类模型包括图表聚类层以及图表分类层;

相应的,所述方法还包括:

将所述各个图表集合中的图表输入所述图表聚类层,获得所述各个图表集合中图表的

聚类结果；

将所述各个图表集合中图表的聚类结果输入所述图表分类层，获得所述各个图表集合中图表的聚类结果对应的分类结果；

基于所述聚类结果以及所述分类结果进行图表筛选，生成三元组图表样本。

7. 根据权利要求6所述的图表识别模型训练方法，其特征在于，还包括：

将所述聚类结果输入预设交叉熵损失函数计算第一损失值；

将所述分类结果输入所述预设交叉熵损失函数计算第二损失值；

判断所述第一损失值以及所述第二损失值是否收敛；

若否，则根据所述第一损失值以及所述第二损失值调整所述图表聚类模型的模型参数，获得更新后的图表聚类模型。

8. 根据权利要求1所述的图表识别模型训练方法，其特征在于，所述根据获得的聚类结果进行图表筛选，生成三元组图表样本，包括：

根据所述聚类结果进行图表筛选，获得多个初始三元组图表样本；

基于每个初始三元组图表样本中各图表样本对应的哈希编码，计算所述每个初始三元组图表样本中目标图表的正样本与所述目标图表间的第一汉明距离，以及，目标图表的负样本与所述目标图表间的第二汉明距离；

在确定所述第一汉明距离与所述第二汉明距离满足约束条件的情况下，将所述初始三元组图表样本确定为所述三元组图表样本。

9. 根据权利要求8所述的图表识别模型训练方法，其特征在于，所述将所述三元组图表样本输入待训练图表识别模型进行训练，获得目标图表识别模型，包括：

将所述三元组图表样本输入至所述待训练图表识别模型进行训练；

对所述三元组图表样本进行训练得到的三元组损失进行平均，得到损失平均值；

根据所述损失平均值调整所述待训练图表识别模型的模型参数直至满足训练条件，获得所述目标图表识别模型。

10. 一种题目批改方法，其特征在于，包括：

接收待识别图片，其中，所述待识别图片中包含待批改题目；

确定所述待批改题目对应的图表检测区域和文本检测区域；

对所述图表检测区域进行图表检测，获得所述图表检测区域中的图表信息，对所述文本检测区域进行文本描述识别获取所述文本检测区域的文本信息；

将所述图表信息输入目标图表识别模型，获得与所述图表信息的匹配度大于第一预设阈值的目标图表，并确定所述目标图表所属题目对应的题目答案；

根据所述题目答案、所述图表信息以及所述文本信息确定所述待批改题目的批改结果；

其中，所述目标图表识别模型是通过权利要求1-9任意一项所述的图表识别模型训练方法训练得到的。

11. 根据权利要求10所述的题目批改方法，其特征在于，还包括：

在未获得与所述图表信息的匹配度大于第一预设阈值的目标图表的情况下，将与所述图表信息的匹配度大于第二预设阈值的图表作为目标图表，并将所述目标图表所属题目对应的题目答案作为参考答案输出。

12. 根据权利要求10所述的题目批改方法,其特征在於,所述确定所述待批改题目对应的图表检测区域和文本检测区域,包括:

将所述待识别图片输入图表检测模型进行目标检测,确定所述待批改题目对应的第一检测区域;

将所述待识别图片输入至文本框检测模型进行目标检测,确定所述待批改题目对应的第二检测区域。

13. 根据权利要求12所述的题目批改方法,其特征在於,所述第一检测区域包括至少一个第一检测子区域,所述第二检测区域包括至少一个第二检测子区域;

相应的,所述根据所述题目答案、所述图表信息以及所述文本信息确定所述待批改题目的批改结果,包括:

根据所述题目答案、每个所述第一检测子区域对应的图表描述信息以及每个所述第二检测子区域对应的文本信息确定待批改题目的批改结果。

14. 一种图表识别模型训练装置,其特征在於,包括:

获取模块,被配置为获取待检测图像集;

检测模块,被配置为对所述待检测图像集中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息;

聚类模块,被配置为根据所述图表类别信息对所述各个图表检测区域中的图表进行筛选,获得不同图表类别对应的图表集合,并对各个图表集合中的图表进行聚类;

训练模块,被配置为根据获得的聚类结果进行图表筛选,生成三元组图表样本,将所述三元组图表样本输入待训练图表识别模型进行训练,获得目标图表识别模型,其中,所述三元组图表样本包括:目标图表、所述目标图表的正样本以及所述目标图表的负样本。

15. 一种题目批改装置,其特征在於,包括:

接收模块,被配置为接收待识别图片,其中,所述待识别图片中包含待批改题目;

确定模块,被配置为确定所述待批改题目对应的图表检测区域和文本检测区域;

检测模块,被配置为对所述图表检测区域进行图表检测,获得所述图表检测区域中的图表信息,对所述文本检测区域进行文本描述识别获取所述文本检测区域的文本信息;

目标图表获取模块,被配置为将所述图表信息输入目标图表识别模型,获得与所述图表信息的匹配度大于第一预设阈值的目标图表,并确定所述目标图表所属题目对应的题目答案;

批改模块,被配置为根据所述题目答案、所述图表信息以及所述文本信息确定所述待批改题目的批改结果;

其中,所述目标图表识别模型是通过权利要求1-9任意一项所述的图表识别模型训练方法训练得到的。

16. 一种计算设备,包括:

存储器和处理器;

所述存储器用于存储计算机可执行指令,所述处理器用于执行所述计算机可执行指令实现权利要求1至9所述的图表识别模型训练方法或权利要求10至13所述的题目批改方法的步骤。

17. 一种计算机可读存储介质,其存储有计算机指令,该指令被处理器执行时实现权利

要求1至9所述的图表识别模型训练方法或权利要求10至13所述的题目批改方法的步骤。

## 图表识别模型训练方法以及装置

### 技术领域

[0001] 本说明书实施例涉及计算机技术领域,特别涉及一种图表识别模型训练方法。本说明书一个或者多个实施例同时涉及一种题目批改方法,一种图表识别模型训练装置,一种题目批改装置,一种计算设备,以及一种计算机可读存储介质。

### 背景技术

[0002] 随着计算机技术的发展,线上教学得到了快速的发展,相应的一些教学工具类产品也应运而生,为学生、老师、家长提供教育辅导上的技术支持和帮助,很多教学工具类产品都能提供将题目拍照进行批改的功能。

[0003] 目前对题目拍照进行批改功能的工具只能解决小学阶段的算式类题型,无法直接处理如珠算等图形类题目的批改,对算式题型的题目批改大多是通过以图搜图的方案来进行替代解决,但是以图搜图的方法只有当题库中有对应的题目的题解时,才能进行题目的批改,而图形类的题型通常数量庞大,类别复杂,当题库中的题目数量较少时,搜索结果不佳,无法得到准确的答案,当题库中的题目数量较多时,通常会带来搜索效率低下,搜索耗时长等问题,因此题库搜索的方式非常依赖题库的指令和以图搜图算法的优劣。

[0004] 因此,如何解决上述问题,提高图形类题目的批改效率,就成为技术人员亟待解决的问题。

### 发明内容

[0005] 有鉴于此,本说明书实施例提供了一种图表识别模型训练方法。本说明书一个或者多个实施例同时涉及一种题目批改方法,一种图表识别模型训练装置,一种题目批改装置,一种计算设备,以及一种计算机可读存储介质,以解决现有技术中存在的技术缺陷。

[0006] 根据本说明书实施例的第一方面,提供了一种图表识别模型训练方法,包括:

[0007] 获取待检测图像集;

[0008] 对所述待检测图像集中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息;

[0009] 根据所述图表类别信息对所述各个图表检测区域中的图表进行筛选,获得不同图表类别对应的图表集合,并对各个图表集合中的图表进行聚类;

[0010] 根据获得的聚类结果进行图表筛选,生成三元组图表样本,将所述三元组图表样本输入待训练图表识别模型进行训练,获得目标图表识别模型,其中,所述三元组图表样本包括:目标图表、所述目标图表的正样本以及所述目标图表的负样本。

[0011] 可选的,所述对各个图表集合中的图表进行聚类之后,所述根据获得的聚类结果进行图表筛选之前,还包括:

[0012] 对所述聚类结果中各聚类类别下的图表进行相似性标注;

[0013] 根据标注结果对所述各聚类类别下的图表进行清洗。

[0014] 可选的,所述对所述待检测图像中的至少一个图表检测区域进行图表检测,获得

各个图表检测区域中图表对应的图表类别信息,包括:

[0015] 将所述待检测图像输入图表检测模型,确定所述待检测图像中的至少一个图表检测区域;

[0016] 对所述至少一个图表检测区域进行图表信息识别,获得所述各个图表检测区域中图表对应的图表类别信息。

[0017] 可选的,所述图表检测模型包括特征提取层、候选区域网络层、空间金字塔池化层、池化层;

[0018] 相应的,所述对所述待检测图像中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息,包括:

[0019] 将所述待检测图像输入至所述特征提取层,提取所述待检测图像的特征图像;

[0020] 将所述特征图像输入至所述候选区域网络层,获取所述特征图像对应的至少一个图表检测区域;

[0021] 将所述特征图像和所述至少一个图表检测区域输入至所述空间金字塔池化层,获得所述至少一个图表检测区域对应的特征向量;

[0022] 将所述特征向量输入至所述池化层进行池化处理,获得所述至少一个图表检测区域对应的图表类别信息。

[0023] 可选的,所述对各个图表集合中的图表进行聚类,包括:

[0024] 将所述各个图表集合中的图表输入图表聚类模型,获得所述各个图表集合中图表的聚类结果。

[0025] 可选的,所述图表聚类模型包括图表聚类层以及图表分类层;

[0026] 相应的,所述方法还包括:

[0027] 将所述各个图表集合中的图表输入所述图表聚类层,获得所述各个图表集合中图表的聚类结果;

[0028] 将所述各个图表集合中图表的聚类结果输入所述图表分类层,获得所述各个图表集合中图表的聚类结果对应的分类结果;

[0029] 基于所述聚类结果以及所述分类结果进行图表筛选,生成三元组图表样本。

[0030] 可选的,所述图表识别模型训练方法,还包括:

[0031] 将所述聚类结果输入预设交叉熵损失函数计算第一损失值;

[0032] 将所述分类结果输入所述预设交叉熵损失函数计算第二损失值;

[0033] 判断所述第一损失值以及所述第二损失值是否收敛;

[0034] 若否,则根据所述第一损失值以及所述第二损失值调整所述图表聚类模型的模型参数,获得更新后的图表聚类模型。

[0035] 可选的,所述根据获得的聚类结果进行图表筛选,生成三元组图表样本,包括:

[0036] 根据所述聚类结果进行图表筛选,获得多个初始三元组图表样本;

[0037] 基于每个初始三元组图表样本中各图表样本对应的哈希编码,计算所述每个初始三元组图表样本中目标图表的正样本与所述目标图表间的第一汉明距离,以及,目标图表的负样本与所述目标图表间的第二汉明距离;

[0038] 在确定所述第一汉明距离与所述第二汉明距离满足约束条件的情况下,将所述初始三元组图表样本确定为所述三元组图表样本。

[0039] 可选的,所述将所述三元组图表样本输入待训练图表识别模型进行训练,获得目标图表识别模型,包括:

[0040] 将所述三元组图表样本输入至所述待训练图表识别模型进行训练;

[0041] 对所述三元组图表样本进行训练得到的三元组损失进行平均,得到损失平均值;

[0042] 根据所述损失平均值调整所述待训练图表识别模型的模型参数直至满足训练条件,获得所述目标图表识别模型。

[0043] 根据本说明书实施例的第二方面,提供了一种题目批改方法,包括:

[0044] 接收待识别图片,其中,所述待识别图片中包含待批改题目;

[0045] 确定所述待批改题目对应的图表检测区域和文本检测区域;

[0046] 对所述图表检测区域进行图表检测,获得所述图表检测区域中的图表信息,对所述文本检测区域进行文本描述识别获取所述文本检测区域的文本信息;

[0047] 将所述图表信息输入目标图表识别模型,获得与所述图表信息的匹配度大于第一预设阈值的目标图表,并确定所述目标图表所属题目对应的题目答案;

[0048] 根据所述题目答案、所述图表信息以及所述文本信息确定所述待批改题目的批改结果;

[0049] 其中,所述目标图表识别模型是通过所述图表识别模型训练方法训练得到的。

[0050] 可选的,所述题目批改方法,还包括:

[0051] 在未获得与所述图表信息的匹配度大于第一预设阈值的目标图表的情况下,将与所述图表信息的匹配度大于第二预设阈值的图表作为目标图表,并将所述目标图表所属题目对应的题目答案作为参考答案输出。

[0052] 可选的,所述确定所述待批改题目对应的图表检测区域和文本检测区域,包括:

[0053] 将所述待识别图片输入图表检测模型进行目标检测,确定所述待批改题目对应的第一检测区域;

[0054] 将所述待识别图片输入至文本框检测模型进行目标检测,确定所述待批改题目对应的第二检测区域。

[0055] 可选的,所述第一检测区域包括至少一个第一检测子区域,所述第二检测区域包括至少一个第二检测子区域;

[0056] 相应的,所述根据所述题目答案、所述图表信息以及所述文本信息确定所述待批改题目的批改结果,包括:

[0057] 根据所述题目答案、每个所述第一检测子区域对应的图表描述信息以及每个所述第二检测子区域对应的文本信息确定待批改题目的批改结果。

[0058] 根据本说明书实施例的第三方面,提供了一种图表识别模型训练装置,包括:

[0059] 获取模块,被配置为获取待检测图像集;

[0060] 检测模块,被配置为对所述待检测图像集中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息;

[0061] 聚类模块,被配置为根据所述图表类别信息对所述各个图表检测区域中的图表进行筛选,获得不同图表类别对应的图表集合,并对各个图表集合中的图表进行聚类;

[0062] 训练模块,被配置为根据获得的聚类结果进行图表筛选,生成三元组图表样本,将所述三元组图表样本输入待训练图表识别模型进行训练,获得目标图表识别模型,其中,所

述三元组图表样本包括：目标图表、所述目标图表的正样本以及所述目标图表的负样本。

[0063] 根据本说明书实施例的第四方面，提供了一种题目批改装置，包括：

[0064] 接收模块，被配置为接收待识别图片，其中，所述待识别图片中包含待批改题目；

[0065] 确定模块，被配置为确定所述待批改题目对应的图表检测区域和文本检测区域；

[0066] 检测模块，被配置为对所述图表检测区域进行图表检测，获得所述图表检测区域中的图表信息，对所述文本检测区域进行文本描述识别获取所述文本检测区域的文本信息；

[0067] 目标图表获取模块，被配置为将所述图表信息输入目标图表识别模型，获得与所述图表信息的匹配度大于第一预设阈值的目标图表，并确定所述目标图表所属题目对应的题目答案；

[0068] 批改模块，被配置为根据所述题目答案、所述图表信息以及所述文本信息确定所述待批改题目的批改结果；

[0069] 其中，所述目标图表识别模型是通过所述图表识别模型训练方法训练得到的。

[0070] 根据本说明书实施例的第五方面，提供了一种计算设备，包括：

[0071] 存储器和处理器；

[0072] 所述存储器用于存储计算机可执行指令，所述处理器用于执行所述计算机可执行指令实现所述图表识别模型训练方法或所述题目批改方法的步骤。

[0073] 根据本说明书实施例的第六方面，提供了一种计算机可读存储介质，其存储有计算机可执行指令，该指令被处理器执行时实现所述图表识别模型训练方法或所述题目批改方法的步骤。

[0074] 本说明书一个实施例通过获取待检测图像集，对所述待检测图像集中的至少一个图表检测区域进行图表检测，获得各个图表检测区域中图表对应的图表类别信息，根据所述图表类别信息对所述各个图表检测区域中的图表进行筛选，获得不同图表类别对应的图表集合，并对各个图表集合中的图表进行聚类，根据获得的聚类结果进行图表筛选，生成三元组图表样本，将所述三元组图表样本输入待训练图表识别模型进行训练，获得目标图表识别模型，其中，所述三元组图表样本包括：目标图表、所述目标图表的正样本以及所述目标图表的负样本；

[0075] 实现了对图表进行检测并聚类后，根据聚类结果进行图表筛选生成三元组图表样本，并利用三元组图表样本进行图表识别模型训练，有利于增加同类图表之间的相似性，同时减小不同类图表之间的相似性，优化了图表识别模型的识别能力，提高了识别结果的准确率以及批改效率。

## 附图说明

[0076] 图1是本说明书一个实施例提供的一种图表识别模型训练方法的处理流程图；

[0077] 图2a是本说明书一个实施例提供的待检测图像的示意图；

[0078] 图2b是本说明书一个实施例提供的图表检测区域的示意图；

[0079] 图3是本说明书一个实施例提供的YoloV3网络的结构图；

[0080] 图4是本说明书一个实施例提供的一种图表聚类的示意图；

[0081] 图5a是本说明书一个实施例提供的一种图表相似性标注的示意图；

- [0082] 图5b是本说明书一个实施例提供的另一种图表相似性标注的示意图；
- [0083] 图6是本说明书一个实施例提供的一种图表聚类模型的聚类框架；
- [0084] 图7是本说明书一个实施例提供的一种三元组样本确定过程的示意图；
- [0085] 图8是本说明书一个实施例提供的一种模型训练过程的示意图；
- [0086] 图9是本说明书另一个实施例提供的一种图表识别模型训练方法的处理过程流程图；
- [0087] 图10是本说明书一个实施例提供的一种题目批改方法的处理流程图；
- [0088] 图11是本说明书一个实施例提供的一种图表识别模型训练装置的示意图；
- [0089] 图12是本说明书一个实施例提供的一种题目批改装置的示意图；
- [0090] 图13是本说明书一个实施例提供的一种计算设备的结构框图。

### 具体实施方式

[0091] 在下面的描述中阐述了很多具体细节以便于充分理解本说明书。但是本说明书能够以很多不同于在此描述的其它方式来实施，本领域技术人员可以在不违背本说明书内涵的情况下做类似推广，因此本说明书不受下面公开的具体实施的限制。

[0092] 在本说明书一个或多个实施例中使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本说明书一个或多个实施例。在本说明书一个或多个实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。还应当理解，本说明书一个或多个实施例中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

[0093] 应当理解，尽管在本说明书一个或多个实施例中可能采用术语第一、第二等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本说明书一个或多个实施例范围的情况下，第一也可以被称为第二，类似地，第二也可以被称为第一。取决于语境，如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

[0094] 首先，对本说明书一个或多个实施例涉及的名词术语进行解释。

[0095] 以图搜图：在指定数据库中搜出相同或相似的图片，一般应用于搜索相同商品或同类商品，在拍照搜题应用领域旨在搜索出相同或相似题目。

[0096] 汉明距离：表示两个哈希编码向量在对应数位上不同的数量，汉明距离小的图像之间相似，汉明距离大的图像之间不相似。

[0097] 局部敏感哈希：是一种哈希方法，使相近的空间点具有相同的哈希值，较远的空间点具有不同的哈希值。

[0098] 三元组损失：一种神经网络损失函数，广泛应用于人脸识别领域，每个三元组由三个样本（假设样本表示为A、B、C）构成，其中A和B样本相似，A和C样本不相似。

[0099] 在本说明书中，提供了一种图表识别模型训练方法，本说明书同时涉及一种题目批改方法，一种图表识别模型训练装置，一种题目批改装置，一种计算设备，以及一种计算机可读存储介质，在下面的实施例中逐一进行详细说明。

[0100] 图1示出了根据本说明书一个实施例提供的一种图表识别模型训练方法的处理流程图，包括步骤102至步骤108。

[0101] 目前对题目拍照进行批改功能的工具只能解决小学阶段的算式类题型,无法直接处理如珠算等图形类题目的批改,对算式题型的题目批改大多是通过以图搜图的方案来进行替代解决,但是以图搜图的方法只有当题库中有对应的题目的题解时,才能进行题目的批改,而图形类的题型通常数量庞大,类别复杂,当题库中的题目数量较少时,搜索结果不佳,无法得到准确的答案,当题库中的题目数量较多时,通常会带来搜索效率低下,搜索耗时长等问题,因此题库搜索的方式非常依赖题库的指令和以图搜图算法的优劣。

[0102] 基于此,本说明书实施例提供一种图表识别模型训练方法,通过获取待检测图像集,对所述待检测图像集中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息,根据所述图表类别信息对所述各个图表检测区域中的图表进行筛选,获得不同图表类别对应的图表集合,并对各个图表集合中的图表进行聚类,根据获得的聚类结果进行图表筛选,生成三元组图表样本,将所述三元组图表样本输入待训练图表识别模型进行训练,获得目标图表识别模型,其中,所述三元组图表样本包括:目标图表、所述目标图表的正样本以及所述目标图表的负样本;

[0103] 实现了对图表进行检测并聚类后,根据聚类结果进行图表筛选生成三元组图表样本,并利用三元组图表样本进行图表识别模型训练,有利于增加同类图表之间的相似性,同时减小不同类图表之间的相似性,优化了图表识别模型的识别能力,提高了识别结果的准确率以及批改效率。

[0104] 步骤102,获取待检测图像集。

[0105] 具体的,待检测图像为用户通过设备终端上传的包括有待批改题目的图像,如用户通过手机拍摄的练习册的照片,还可以为用户通过应用软件扫描练习册得到的图片等等,在本申请中,对待检测图像的获取方式不做限制。

[0106] 待检测图像中包括至少一道待批改题目,其中,待批改题目为图形类题型,如珠算题、识图题等等。

[0107] 在本说明书提供的一具体实施方式中,以待检测图像为用户通过手机拍摄的练习册照片为例,待检测图像中的待批改题目为图形类的识图题,通过观察图表中涂色部分的个数来确定数字。

[0108] 步骤104,对所述待检测图像集中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息。

[0109] 具体的,待检测图像集中的一张待检测图像的示意图如图2a所示。

[0110] 实际应用中,图表类型有十多种,包括木棍、地图、阴影图形、钟表、图形组等,对图2a所示的待检测图像进行图表检测区域识别,获取所述待检测图像对应的多个图表检测区域,图表检测区域的示意图参见图2b,对图2b中的图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息,并根据所述图表类别信息确定图2b中各个图表检测区域中图表的图表类型均为阴影图形。

[0111] 具体实施时,对所述待检测图像中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息,具体可通过以下方式实现:

[0112] 将所述待检测图像输入图表检测模型,确定所述待检测图像中的至少一个图表检测区域;

[0113] 对所述至少一个图表检测区域进行图表信息识别,获得所述各个图表检测区域中

图表对应的图表类别信息。

[0114] 具体的,图表检测模型用于对待检测图像中的图表进行检测,即给定一个待检测图像,找到图像中每个图表的位置,并确定相应位置对应的图表的类别,图表检测模型通常是在一组固定的训练集上进行训练的,图表检测模型需要确定图像中目标图表的位置信息和对目标图表进行分类,通过神经网络模型的方式实现检测目的,有效地提升了识别准确率。

[0115] 实际应用中,所述图表检测模型可以为YoloV3 (Yolo, You Only Look Once,) 模型,Yolo是一种通用的目标检测模型,基于简单的端到端网络,直接采用回归方法对目标对象所在的框进行检测以及分类。YoloV3在Yolo模型基础上采用多尺度特征进行目标检测。

[0116] YoloV3模型使用的骨干网络为darknet-53网络,darknet-53网络的网络结构中的前52层,如下图3所示,图3示出了YoloV3网络的结构图,其中,DBL是YoloV3的基本组件,为卷积+BN+Leaky relu,对于YoloV3来说,BN和Leaky relu已经是和卷积层不可分离的部分,共同构成了最小组件。

[0117] resn中的n代表数字,有res1,res2,⋯,res8等等,表示这个res\_block里含有多少个res\_unit。是YoloV3的大组件,YoloV3开始借鉴了ResNet的残差结构,使用这种结构可以让网络结构更深。

[0118] Concat为张量拼接,将darknet中间层和后面的某一层的上采样进行拼接。拼接的操作和残差层add的操作是不一样的,拼接会扩充张量的维度,而add只是直接相加不会导致张量维度的改变。

[0119] 如图3所示,YoloV3输出3个不同尺度的特征图像Y1、Y2、Y3,其中Y1、Y2、Y3的深度都为255,边长规律为13:26:52,在每个特征图像中会输出3个,共计9个预测框,再从9个预测框中找到目标存在可能性得分最高的预测框作为所述待批改题目对应的第一检测区域。

[0120] 具体实施时,所述图表检测模型包括特征提取层、候选区域网络层、空间金字塔池化层、池化层;

[0121] 相应的,所述对所述待检测图像中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息,包括:

[0122] 将所述待检测图像输入至所述特征提取层,提取所述待检测图像的特征图像;

[0123] 将所述特征图像输入至所述候选区域网络层,获取所述特征图像对应的至少一个图表检测区域;

[0124] 将所述特征图像和所述至少一个图表检测区域输入至所述空间金字塔池化层,获得所述至少一个图表检测区域对应的特征向量;

[0125] 将所述特征向量输入至所述池化层进行池化处理,获得所述至少一个图表检测区域对应的图表类别信息。

[0126] 具体的,特征提取层用于提取待检测图像的特征图像,特征提取层可以为移除了全连接层的残差结构的卷积神经网络,残差结构的卷积神经网络可以为ResNet、ResNet18、ResNet50等。

[0127] 在本申请提供的实施例中,将图2a所示的待检测图像输入至移除了全连接层的ResNet50卷积神经网络,ResNet50卷积神经网络包括第一卷积层组、第二卷积层组、第三卷积层组、第四卷积层组和第五卷积层组,提取所述待检测图像的特征图像。

[0128] 候选区域网络层(RPN)也是全卷积网络(FCN,fully-convolutional network),可以针对生成检测建议框的任务端到端的进行训练,能够同时预测出object的边界和分数,只是在CNN上额外增加了2个卷积层(全卷积层cls和reg)。RPN网络的输入可以是任意大小的图片。将特征图像输入至RPN中,RPN将每个特征图像的位置编码生成一个特征向量,对每一个位置输出一个objectness score和regressedbounds for k个候选区域(region proposal),即在每个卷积映射位置输出这个位置上多种尺度(3种)和长宽比(3种)的k个(3\*3=9)区域建议的物体得分和候选区域信息,通过RPN层获取所述特征图像对应的多个子候选区域(图表检测区域)。

[0129] 在本申请提供的实施例中,沿用上例,将所述特征图像输入至所述候选区域网络层,获取所述特征图像对应的多个子候选区域(图表检测区域),参见图2b,图2b示出了候选区域网络输出的所述特征图像对应的多个子候选区域(图表检测区域)。

[0130] 金字塔池化层(Spatial Pyramid Pooling,SPP)可以将不同大小的特征图片作为输入,将一个特征图片从不同角度进行特征提取再聚合,生成固定大小的输出。从多尺度特征中提取出固定大小的特征向量,解决了特征图像大小不一的问题。通过特征向量和多个子候选区域获得每个子候选区域对应固定大小的候选特征向量。

[0131] 在本申请提供的实施例中,将候选区域网络输出的所述特征图像和对应的多个子候选区域输入至空间金字塔池化层,获得每个子候选区域对应的候选特征向量。

[0132] 将每个子候选区域对应的候选特征向量进行池化,获得所述待检测图像的至少一个图表检测区域对应的图表类别信息。

[0133] 在本申请提供的实施例中,将每个所述子候选区域和所述特征图像输入至空间金字塔池化层,获得每个子候选区域对应的固定大小的候选特征向量。从多尺度特征中提取出固定大小的特征向量,解决了特征图像大小不一的问题。

[0134] 步骤106,根据所述图表类别信息对所述各个图表检测区域中的图表进行筛选,获得不同图表类别对应的图表集合,并对各个图表集合中的图表进行聚类。

[0135] 具体的,如前所述,带图大题中含有的图表类型有十多种,包括木棍、地图、阴影图形、钟表、图形组等,在获得各个图表检测区域中图表对应的图表类别信息后,可根据所述图表类别信息确定图表类型,并根据图表类型对图表进行筛选,将属于同一图表类型的图表进行筛选生成对应的图表集合,例如,将阴影图形进行筛选,生成阴影图形对应的图表集合,将钟表类型的图表进行筛选,生成钟表对应的图表集合等。

[0136] 获取不同图表类别对应的图表集合后,可对各个图表集合中的图表进行聚类,即将同一图表类别中相同的图表聚为一类。

[0137] 以阴影图形为例,图表聚类的示意图如图4所示,由于口算产品的带图大题里含有大量的阴影图形,并且每个阴影图形的形状、颜色可能均存在差异,如图4所示,第一行的两个阴影图形相同,第二行的两个阴影图形不同,因此,将第一行的两个阴影图形聚为一类。

[0138] 本说明书实施例采用端到端的深度聚类方法,提升了图表聚类质量,减小了数据标注成本。

[0139] 另外,对各个图表集合中的图表进行聚类之后,根据获得的聚类结果进行图表筛选之前,还可对聚类结果中的图表进行相似性标注,具体可通过以下方式实现:

[0140] 对所述聚类结果中各聚类类别下的图表进行相似性标注;

[0141] 根据标注结果对所述各聚类类别下的图表进行清洗。

[0142] 具体的,在进行图表聚类之后,得到多个聚类类别,同一个类别的阴影图形相似,不同类的阴影图形不相似。对同一类别下的图表进行相似性标注的目的是为了清除某个聚类类别内的脏数据,所谓脏数据是在一个聚类类别内不同于其他图表的数据。清除脏数据后,可保证同一个聚类类别内的图表都是相似的。

[0143] 进行相似性标注的标签分为三种:“1.相同”、“2.不相同但非常相似”、“3.不同”。一种图表相似性标注的示意图如图5a,两个阴影图形的形状,长度都相同,需要标注为“1.相同”。

[0144] 另一种图表相似性标注的示意图如图5b,第一个阴影图形由6个阴影三角形构成,第二个阴影图形由3个阴影三角形构成,因此需标注为不同。

[0145] 通过进行数据清洗,有利于提高模型的训练质量。

[0146] 此外,还可通过图表聚类模型对图表集合中的各个图表进行聚类,即将所述各个图表集合中的图表输入图表聚类模型,获得所述各个图表集合中图表的聚类结果。

[0147] 具体实施时,所述图表聚类模型包括图表聚类层以及图表分类层;将所述各个图表集合中的图表输入所述图表聚类层,获得所述各个图表集合中图表的聚类结果;将所述各个图表集合中图表的聚类结果输入所述图表分类层,获得所述各个图表集合中图表的聚类结果对应的分类结果。

[0148] 进一步的,还可将所述聚类结果输入预设交叉熵损失函数计算第一损失值,将所述分类结果输入所述预设交叉熵损失函数计算第二损失值,判断所述第一损失值以及所述第二损失值是否收敛,若否,则根据所述第一损失值以及所述第二损失值调整所述图表聚类模型的模型参数,获得更新后的图表聚类模型。

[0149] 具体的,图表聚类模型的聚类框架如图6,训练数据是通过图表检测切分出来的海量数据,主干网络是一种端到端训练的卷积神经网络,第一层为输入层(input layer),第二层为图表聚类层,第三层为图表分类层,第四层为输出层(output layer),卷积神经网络提取的特征向量是用于聚类的原始数据,聚类算法采用传统的k均值聚类,聚类的类别作为样本的伪标签,从而构成了弱监督的标签分类任务,图表聚类模型的整体损失函数由两部分组成,分类损失和聚类损失,融合分类和聚类损失对所述图表聚类模型进行迭代式训练。

[0150] 本说明书实施例采用端到端的深度卷积网络,将聚类类别作为伪标签,聚类和分类两个模块迭代式训练,提升了聚类的准确度,并且采用端到端的深度聚类方法,提升了图像聚类质量,减小了数据标注成本,并有利于提升模型训练结果的准确性。

[0151] 步骤108,根据获得的聚类结果进行图表筛选,生成三元组图表样本,将所述三元组图表样本输入待训练图表识别模型进行训练,获得目标图表识别模型,其中,所述三元组图表样本包括:目标图表、所述目标图表的正样本以及所述目标图表的负样本。

[0152] 具体的,所述图表识别模型可以是深度哈希模型,深度哈希模型框架的第一步是从原始的聚类结果中选择合适的三元组,三元组由三个样本组成,分别是目标样本、目标样本的正样本、目标样本的负样本,其中,目标样本是随机选择出来的样本,目标样本的正样本跟目标样本在同一个聚类类别内,属于与目标样本相同的图表。目标样本的负样本跟目标样本以及目标样本的正样本均不在一个聚类类别内,属于与目标样本不同的图表。

[0153] 三元组样本确定过程的示意图如图7所示,对各个图表集合中的图表进行聚类获

得4个聚类类别,分别为cluster1、cluster2、cluster3、cluster4,从4个聚类类别中筛选获得N个三元组样本,分别为Group 1至Group N,获得三元组样本后,将所述三元组图表样本输入待训练图表识别模型进行训练,获得目标图表识别模型。

[0154] 另外,所述根据获得的聚类结果进行图表筛选,生成三元组图表样本,包括:

[0155] 根据所述聚类结果进行图表筛选,获得多个初始三元组图表样本;

[0156] 基于每个初始三元组图表样本中各图表样本对应的哈希编码,计算所述每个初始三元组图表样本中目标图表的正样本与所述目标图表间的第一汉明距离,以及,目标图表的负样本与所述目标图表间的第二汉明距离;

[0157] 在确定所述第一汉明距离与所述第二汉明距离满足约束条件的情况下,将所述初始三元组图表样本确定为所述三元组图表样本。

[0158] 具体的,选择三元组图表样本时,需要计算样本中图表之间的汉明距离,假设d1是目标样本与目标样本的正样本之间的汉明距离,d2是目标样本与目标样本的负样本之间的汉明距离。如果不考虑其他因素,三元组图表样本的数量将呈指数级增长,且深度学习模型不易收敛。需要增加约束条件,使距离满足 $d1 + margin > d2$ ,因此,在筛选获得多个初始三元组图表样本,并确定初始三元组图标样本中的第一汉明距离d1与所述第二汉明距离d2满足上述约束条件的情况下,将所述初始三元组图表样本确定为所述三元组图表样本。

[0159] 深度哈希模型采用三元组的损失函数,在选择合适的三元组时避免了模型不收敛的问题,有利于保证模型训练结果的准确性。

[0160] 此外,将所述各个图表集合中的图表输入所述图表聚类层,获得所述各个图表集合中图表的聚类结果,并将所述各个图表集合中图表的聚类结果输入所述图表分类层,获得所述各个图表集合中图表的聚类结果对应的分类结果后,可基于所述聚类结果以及所述分类结果进行图表筛选,生成三元组图表样本。

[0161] 进一步的,所述将所述三元组图表样本输入待训练图表识别模型进行训练,获得目标图表识别模型,包括:

[0162] 将所述三元组图表样本输入至所述待训练图表识别模型进行训练;

[0163] 对所述三元组图表样本进行训练得到的三元组损失进行平均,得到损失平均值;

[0164] 根据所述损失平均值调整所述待训练图表识别模型的模型参数直至满足训练条件,获得所述目标图表识别模型。

[0165] 具体的,选择三元组图表样本后,采用端到端的卷积神经网络训练,卷积神经网络可以采用resnet50,用来做图像的特征提取。具体的模型训练过程的示意图如图8所示,ResNet50卷积神经网络包括第一卷积层组、第二卷积层组、第三卷积层组、第四卷积层组和第五卷积层组,将三元组图表样本输入所述ResNet50卷积神经网络,提取所述待检测图像的特征,并获取对所述三元组图表样本进行训练得到的三元组损失进行平均,得到损失平均值,根据所述损失平均值调整所述待训练图表识别模型的模型参数直至满足训练条件,获得所述目标图表识别模型。

[0166] 由于ResNet50卷积神经网络输出层的激活函数为tanh函数,将输出值的范围约束在-1到1之间,输出128维度的浮点向量,随着ResNet50卷积神经网络训练的梯度下降,相似样本之间的距离越来越近,不相似样本之间的距离越来越远,从而ResNet50卷积神经网络能输出紧致的二值哈希编码。

[0167] 深度哈希模型采用三元组损失函数,在选择合适的三元组时避免了模型不收敛的问题,另外,由于存储128维度的浮点向量需要消耗大量的存储空间,为了减小题库的存储空间,本说明书实施例采用局部敏感哈希方法(LSH)将128维度的浮点向量压缩为64维的二值向量,同时保留样本之间的相对相似性信息。如果两个浮点向量在原有的128维数据空间是相似的,那么分别经过LSH哈希函数转换以后,它们在64维数据空间中也具有很高的相似度。反之,如果它们在转换前是不相似的,那么经过转换后它们仍不具有相似性。此外,LSH还能在一定程度上减小以图搜图的搜索时长。

[0168] 因此,采用局部敏感哈希方法对模型生成的浮点向量转化为二值向量,节省了哈希编码的存储空间,同时便于建立数据库索引,提升了以图搜图的检索速度。

[0169] 本说明书实施例通过获取待检测图像集,对所述待检测图像集中的至少一个图表检测区域进行图表检测,获得各个图表检测区域中图表对应的图表类别信息,根据所述图表类别信息对所述各个图表检测区域中的图表进行筛选,获得不同图表类别对应的图表集合,并对各个图表集合中的图表进行聚类,根据获得的聚类结果进行图表筛选,生成三元组图表样本,将所述三元组图表样本输入待训练图表识别模型进行训练,获得目标图表识别模型,其中,所述三元组图表样本包括:目标图表、所述目标图表的正样本以及所述目标图表的负样本;

[0170] 实现了对图表进行检测并聚类后,根据聚类结果进行图表筛选生成三元组图表样本,并利用三元组图表样本进行图表识别模型训练,有利于增加同类图表之间的相似性,同时减小不同类图表之间的相似性,优化了图表识别模型的识别能力,提高了识别结果的准确率以及批改效率。

[0171] 下述结合附图9,以本说明书提供的另一种图表识别模型训练方法的处理流程图,具体步骤包括步骤902至步骤926。

[0172] 步骤902,获取待检测图像集。

[0173] 步骤904,将所述待检测图像输入图表检测模型,确定所述待检测图像中的至少一个图表检测区域。

[0174] 步骤906,对所述至少一个图表检测区域进行图表信息识别,获得所述各个图表检测区域中图表对应的图表类别信息。

[0175] 步骤908,根据所述图表类别信息对所述各个图表检测区域中的图表进行筛选,获得不同图表类别对应的图表集合。

[0176] 步骤910,将所述各个图表集合中的图表输入图表聚类模型,获得所述各个图表集合中图表的聚类结果。

[0177] 步骤912,对所述聚类结果中各聚类类别下的图表进行相似性标注。

[0178] 步骤914,根据标注结果对所述各聚类类别下的图表进行清洗,获得目标聚类结果。

[0179] 步骤916,根据所述目标聚类结果进行图表筛选,获得多个初始三元组图表样本。

[0180] 步骤918,基于每个初始三元组图表样本中各图表样本对应的哈希编码,计算所述每个初始三元组图表样本中目标图表的正样本与所述目标图表间的第一汉明距离,以及,目标图表的负样本与所述目标图表间的第二汉明距离。

[0181] 步骤920,在确定所述第一汉明距离与所述第二汉明距离满足约束条件的情况下,

将所述初始三元组图表样本确定为所述三元组图表样本。

[0182] 具体的,所述三元组图表样本包括:目标图表、所述目标图表的正样本以及所述目标图表的负样本。

[0183] 步骤922,将所述三元组图表样本输入至所述待训练图表识别模型进行训练。

[0184] 步骤924,对所述三元组图表样本进行训练得到的三元组损失进行平均,得到损失平均值。

[0185] 步骤926,根据所述损失平均值调整所述待训练图表识别模型的模型参数直至满足训练条件,获得所述目标图表识别模型。

[0186] 本说明书实施例对图表进行检测并聚类后,根据聚类结果进行图表筛选生成三元组图表样本,并利用三元组图表样本进行图表识别模型训练,有利于增加同类图表之间的相似性,同时减小不同类图表之间的相似性,优化了图表识别模型的识别能力,提高了识别结果的准确率。

[0187] 图10示出了根据本说明书一个实施例提供的一种题目批改方法的处理流程图,包括步骤1002至步骤1010。

[0188] 步骤1002,接收待识别图片,其中,所述待识别图片中包含待批改题目。

[0189] 步骤1004,确定所述待批改题目对应的图表检测区域和文本检测区域。

[0190] 步骤1006,对所述图表检测区域进行图表检测,获得所述图表检测区域中的图表信息,对所述文本检测区域进行文本描述识别获取所述文本检测区域的文本信息。

[0191] 步骤1008,将所述图表信息输入目标图表识别模型,获得与所述图表信息的匹配度大于第一预设阈值的目标图表,并确定所述目标图表所属题目对应的题目答案。

[0192] 步骤1010,根据所述题目答案、所述图表信息以及所述文本信息确定所述待批改题目的批改结果;其中,所述目标图表识别模型是通过所述图表识别模型训练方法训练得到的。

[0193] 具体的,所述待识别图片中包括至少一道待批改题目,其中,待批改题目为图形类题型,如珠算题、识图题等等,用户需要对待批改题目的答案进行批改,判断答题是否正确。

[0194] 另外,所述目标图表识别模型通过前述图表识别模型训练方法训练获得,通过所述目标图表识别模型识别获得与待批改题目中的图表相同或相似(匹配度大于第一预设阈值)的目标图表,并根据所述目标图表所属题目对应的题目答案对所述待批改题目进行题目批改。

[0195] 可选地,在未获得与所述图表信息的匹配度大于第一预设阈值的目标图表的情况下,将与所述图表信息的匹配度大于第二预设阈值的图表作为目标图表,并将所述目标图表所属题目对应的题目答案作为参考答案输出。

[0196] 具体的,在未获得与待批改题目中的图表相同或相似(匹配度大于第一预设阈值)的目标图表的情况下,可将与待批改题目中的图表的匹配度大于第二预设阈值的图表作为目标图表,并将所述目标图表所属题目对应的题目答案作为参考答案输出,其中,所述第一预设阈值大于所述第二预设阈值,预设阈值与匹配度成正比,预设阈值越大,则待批改题目中的图表与目标图表间的匹配度越高。

[0197] 可选地,所述确定所述待批改题目对应的图表检测区域和文本检测区域,包括:

[0198] 将所述待识别图片输入图表检测模型进行目标检测,确定所述待批改题目对应的

第一检测区域；

[0199] 将所述待识别图片输入至文本框检测模型进行目标检测，确定所述待批改题目对应的第二检测区域。

[0200] 进一步的，所述第一检测区域包括至少一个第一检测子区域，所述第二检测区域包括至少一个第二检测子区域；

[0201] 相应的，所述根据所述题目答案、所述图表信息以及所述文本信息确定所述待批改题目的批改结果，包括：

[0202] 根据所述题目答案、每个所述第一检测子区域对应的图表描述信息以及每个所述第二检测子区域对应的文本信息确定待批改题目的批改结果。

[0203] 具体的，通过所述目标图表识别模型识别获得与待批改题目中的图表相同或相似（匹配度大于第一预设阈值）的目标图表，并根据所述目标图表所属题目对应的题目答案对所述待批改题目进行题目批改，在题目批改过程中，还需结合待批改题目的题目信息、用户填写的答案信息以及目标图表所属题目对应的题目信息。

[0204] 本说明书实施例通过接收待识别图片，确定所述待批改题目对应的图表检测区域和文本检测区域，对所述图表检测区域进行图表检测，获得所述图表检测区域中的图表信息，对所述文本检测区域进行文本描述识别获取所述文本检测区域的文本信息，将所述图表信息输入目标图表识别模型，获得与所述图表信息的匹配度大于第一预设阈值的目标图表，并确定所述目标图表所属题目对应的题目答案，根据所述题目答案、所述图表信息以及所述文本信息确定所述待批改题目的批改结果。

[0205] 实现了根据所述图表信息和所述文本信息确定待批改题目的批改结果，通过检测识别第一检测区域的图表信息和识别第二检测区域的文本信息，将两者与目标图表所属题目以及题目对应的题目答案进行比对，以进行题目批改，利用以图搜图的方式，从题库中搜索相似图片，并结合题目中的文本信息，对题库中的题目进行综合排序，对于命中的题目进行批改，对于没有命中的题目提供相似题目的解析，有利于保证题目批改结果的准确率，同时有利于提升用户的服务体验。

[0206] 与上述方法实施例相对应，本说明书还提供了图表识别模型训练装置实施例，图11示出了本说明书一个实施例提供的一种图表识别模型训练装置的示意图。如图11所示，该装置包括：

[0207] 获取模块1102，被配置为获取待检测图像集；

[0208] 检测模块1104，被配置为对所述待检测图像集中的至少一个图表检测区域进行图表检测，获得各个图表检测区域中图表对应的图表类别信息；

[0209] 聚类模块1106，被配置为根据所述图表类别信息对所述各个图表检测区域中的图表进行筛选，获得不同图表类别对应的图表集合，并对各个图表集合中的图表进行聚类；

[0210] 训练模块1108，被配置为根据获得的聚类结果进行图表筛选，生成三元组图表样本，将所述三元组图表样本输入待训练图表识别模型进行训练，获得目标图表识别模型，其中，所述三元组图表样本包括：目标图表、所述目标图表的正样本以及所述目标图表的负样本。

[0211] 可选地，所述图表识别模型训练装置，还包括：

[0212] 标注模块，被配置为对所述聚类结果中各聚类类别下的图表进行相似性标注；

- [0213] 清洗模块,被配置为根据标注结果对所述各聚类类别下的图表进行清洗。
- [0214] 可选地,所述检测模块1102,包括:
- [0215] 确定子模块,被配置为将所述待检测图像输入图表检测模型,确定所述待检测图像中的至少一个图表检测区域;
- [0216] 识别子模块,被配置为对所述至少一个图表检测区域进行图表信息识别,获得所述各个图表检测区域中图表对应的图表类别信息。
- [0217] 可选地,所述图表检测模型包括特征提取层、候选区域网络层、空间金字塔池化层、池化层;
- [0218] 相应的,所述检测模块1102,包括:
- [0219] 特征图像提取子模块,被配置为将所述待检测图像输入至所述特征提取层,提取所述待检测图像的特征图像;
- [0220] 图表检测区域获取子模块,被配置为将所述特征图像输入至所述候选区域网络层,获取所述特征图像对应的至少一个图表检测区域;
- [0221] 特征向量获取子模块,被配置为将所述特征图像和所述至少一个图表检测区域输入至所述空间金字塔池化层,获得所述至少一个图表检测区域对应的特征向量;
- [0222] 图表类别信息获取子模块,被配置为将所述特征向量输入至所述池化层进行池化处理,获得所述至少一个图表检测区域对应的图表类别信息。
- [0223] 可选地,所述聚类模块1106,包括:
- [0224] 聚类子模块,被配置为将所述各个图表集合中的图表输入图表聚类模型,获得所述各个图表集合中图表的聚类结果。
- [0225] 可选地,所述图表聚类模型包括图表聚类层以及图表分类层;
- [0226] 相应的,所述图表识别模型训练装置,还包括:
- [0227] 聚类结果获取模块,被配置为将所述各个图表集合中的图表输入所述图表聚类层,获得所述各个图表集合中图表的聚类结果;
- [0228] 分类结果获取模块,被配置为将所述各个图表集合中图表的聚类结果输入所述图表分类层,获得所述各个图表集合中图表的聚类结果对应的分类结果;
- [0229] 样本生成模块,被配置为基于所述聚类结果以及所述分类结果进行图表筛选,生成三元组图表样本。
- [0230] 可选地,所述图表识别模型训练装置,还包括:
- [0231] 第一计算模块,被配置为将所述聚类结果输入预设交叉熵损失函数计算第一损失值;
- [0232] 第二计算模块,被配置为将所述分类结果输入所述预设交叉熵损失函数计算第二损失值;
- [0233] 判断模块,被配置为判断所述第一损失值以及所述第二损失值是否收敛;
- [0234] 若所述判断模块的运行结果为否,则运行调整模块;
- [0235] 所述调整模块,被配置为根据所述第一损失值以及所述第二损失值调整所述图表聚类模型的模型参数,获得更新后的图表聚类模型。
- [0236] 可选地,所述训练模块1108,包括:
- [0237] 筛选子模块,被配置为根据所述聚类结果进行图表筛选,获得多个初始三元组图

表样本；

[0238] 汉明距离计算子模块，被配置为基于每个初始三元组图表样本中各图表样本对应的哈希编码，计算所述每个初始三元组图表样本中目标图表的正样本与所述目标图表间的第一汉明距离，以及，目标图表的负样本与所述目标图表间的第二汉明距离；

[0239] 样本确定子模块，被配置为在确定所述第一汉明距离与所述第二汉明距离满足约束条件的情况下，将所述初始三元组图表样本确定为所述三元组图表样本。

[0240] 可选地，所述训练模块1108，包括：

[0241] 训练子模块，被配置为将所述三元组图表样本输入至所述待训练图表识别模型进行训练；

[0242] 损失平均值计算模块，被配置为对所述三元组图表样本进行训练得到的三元组损失进行平均，得到损失平均值；

[0243] 参数调整子模块，被配置为根据所述损失平均值调整所述待训练图表识别模型的模型参数直至满足训练条件，获得所述目标图表识别模型。

[0244] 上述为本实施例的一种图表识别模型训练装置的示意性方案。需要说明的是，该图表识别模型训练装置的技术方案与上述的图表识别模型训练方法的技术方案属于同一构思，图表识别模型训练装置的技术方案未详细描述的细节内容，均可以参见上述图表识别模型训练方法的技术方案的描述。

[0245] 与上述方法实施例相对应，本说明书还提供了题目批改装置实施例，图12示出了本说明书一个实施例提供的一种题目批改装置的示意图。如图12所示，该装置包括：

[0246] 接收模块1202，被配置为接收待识别图片，其中，所述待识别图片中包含待批改题目；

[0247] 确定模块1204，被配置为确定所述待批改题目对应的图表检测区域和文本检测区域；

[0248] 检测模块1206，被配置为对所述图表检测区域进行图表检测，获得所述图表检测区域中的图表信息，对所述文本检测区域进行文本描述识别获取所述文本检测区域的文本信息；

[0249] 目标图表获取模块1208，被配置为将所述图表信息输入目标图表识别模型，获得与所述图表信息的匹配度大于第一预设阈值的目标图表，并确定所述目标图表所属题目对应的题目答案；

[0250] 批改模块1210，被配置为根据所述题目答案、所述图表信息以及所述文本信息确定所述待批改题目的批改结果；其中，所述目标图表识别模型是通过所述图表识别模型训练方法训练得到的。

[0251] 可选地，所述题目批改装置，还包括：

[0252] 参考答案输出模块，被配置为在未获得与所述图表信息的匹配度大于第一预设阈值的目标图表的情况下，将与所述图表信息的匹配度大于第二预设阈值的图表作为目标图表，并将所述目标图表所属题目对应的题目答案作为参考答案输出。

[0253] 可选地，所述确定模块1204，包括：

[0254] 第一确定模块，被配置为将所述待识别图片输入图表检测模型进行目标检测，确定所述待批改题目对应的第一检测区域；

[0255] 第二确定模块,被配置为将所述待识别图片输入至文本框检测模型进行目标检测,确定所述待批改题目对应的第二检测区域。

[0256] 可选地,所述第一检测区域包括至少一个第一检测子区域,所述第二检测区域包括至少一个第二检测子区域;

[0257] 相应的,所述批改模块1210,包括:

[0258] 批改结果确定子模块,被配置为根据所述题目答案、每个所述第一检测子区域对应的图表描述信息以及每个所述第二检测子区域对应的文本信息确定待批改题目的批改结果。

[0259] 上述为本实施例的一种题目批改装置的示意性方案。需要说明的是,该题目批改装置的技术方案与上述的题目批改方法的技术方案属于同一构思,题目批改装置的技术方案未详细描述的细节内容,均可以参见上述题目批改方法的技术方案的描述。

[0260] 图13示出了根据本说明书一个实施例提供的一种计算设备1300的结构框图。该计算设备1300的部件包括但不限于存储器1310和处理器1320。处理器1320与存储器1310通过总线1330相连接,数据库1350用于保存数据。

[0261] 计算设备1300还包括接入设备1340,接入设备1340使得计算设备1300能够经由一个或多个网络1360通信。这些网络的示例包括公用交换电话网(PSTN)、局域网(LAN)、广域网(WAN)、个域网(PAN)或诸如因特网的通信网络的组合。接入设备1340可以包括有线或无线的任何类型的网络接口(例如,网络接口卡(NIC))中的一个或多个,诸如IEEE802.11无线局域网(WLAN)无线接口、全球微波互联接入(Wi-MAX)接口、以太网接口、通用串行总线(USB)接口、蜂窝网络接口、蓝牙接口、近场通信(NFC)接口,等等。

[0262] 在本说明书的一个实施例中,计算设备1300的上述部件以及图13中未示出的其他部件也可以彼此相连接,例如通过总线。应当理解,图13所示的计算设备结构框图仅仅是出于示例的目的,而不是对本说明书范围的限制。本领域技术人员可以根据需要,增添或替换其他部件。

[0263] 计算设备1300可以是任何类型的静止或移动计算设备,包括移动计算机或移动计算设备(例如,平板计算机、个人数字助理、膝上型计算机、笔记本计算机、上网本等)、移动电话(例如,智能手机)、可佩戴的计算设备(例如,智能手表、智能眼镜等)或其他类型的移动设备,或者诸如台式计算机或PC的静止计算设备。计算设备1300还可以是移动式或静止式的服务器。

[0264] 其中,所述存储器1310用于存储计算机可执行指令,处理器1320用于执行计算机可执行指令实现所述图表识别模型训练方法或所述题目批改方法的步骤。

[0265] 上述为本实施例的一种计算设备的示意性方案。需要说明的是,该计算设备的技术方案与上述的图表识别模型训练方法或所述题目批改方法的技术方案属于同一构思,计算设备的技术方案未详细描述的细节内容,均可以参见上述图表识别模型训练方法或所述题目批改方法的技术方案的描述。

[0266] 本说明书一实施例还提供一种计算机可读存储介质,其存储有计算机指令,该指令被处理器执行时以用于实现所述图表识别模型训练方法或所述题目批改方法的步骤。

[0267] 上述为本实施例的一种计算机可读存储介质的示意性方案。需要说明的是,该存储介质的技术方案与上述的图表识别模型训练方法或所述题目批改方法的技术方案属于

同一构思,存储介质的技术方案未详细描述的细节内容,均可以参见上述图表识别模型训练方法或所述题目批改方法的技术方案的描述。

[0268] 上述对本说明书特定实施例进行了描述。其它实施例在所附权利要求书的范围内。在一些情况下,在权利要求书中记载的动作或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外,在附图中描绘的过程不一定要求示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中,多任务处理和并行处理也是可以的或者可能是有利的。

[0269] 所述计算机指令包括计算机程序代码,所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括:能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、电载波信号、电信信号以及软件分发介质等。需要说明的是,所述计算机可读介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减,例如在某些司法管辖区,根据立法和专利实践,计算机可读介质不包括电载波信号和电信信号。

[0270] 需要说明的是,对于前述的各方法实施例,为了简便描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本说明书实施例并不受所描述的动作顺序的限制,因为依据本说明书实施例,某些步骤可以采用其它顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定都是本说明书实施例所必须的。

[0271] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述的部分,可以参见其它实施例的相关描述。

[0272] 以上公开的本说明书优选实施例只是用于帮助阐述本说明书。可选实施例并没有详尽叙述所有的细节,也不限制该发明仅为所述的具体实施方式。显然,根据本说明书实施例的内容,可作很多的修改和变化。本说明书选取并具体描述这些实施例,是为了更好地解释本说明书实施例的原理和实际应用,从而使所属技术领域技术人员能很好地理解和利用本说明书。本说明书仅受权利要求书及其全部范围和等效物的限制。

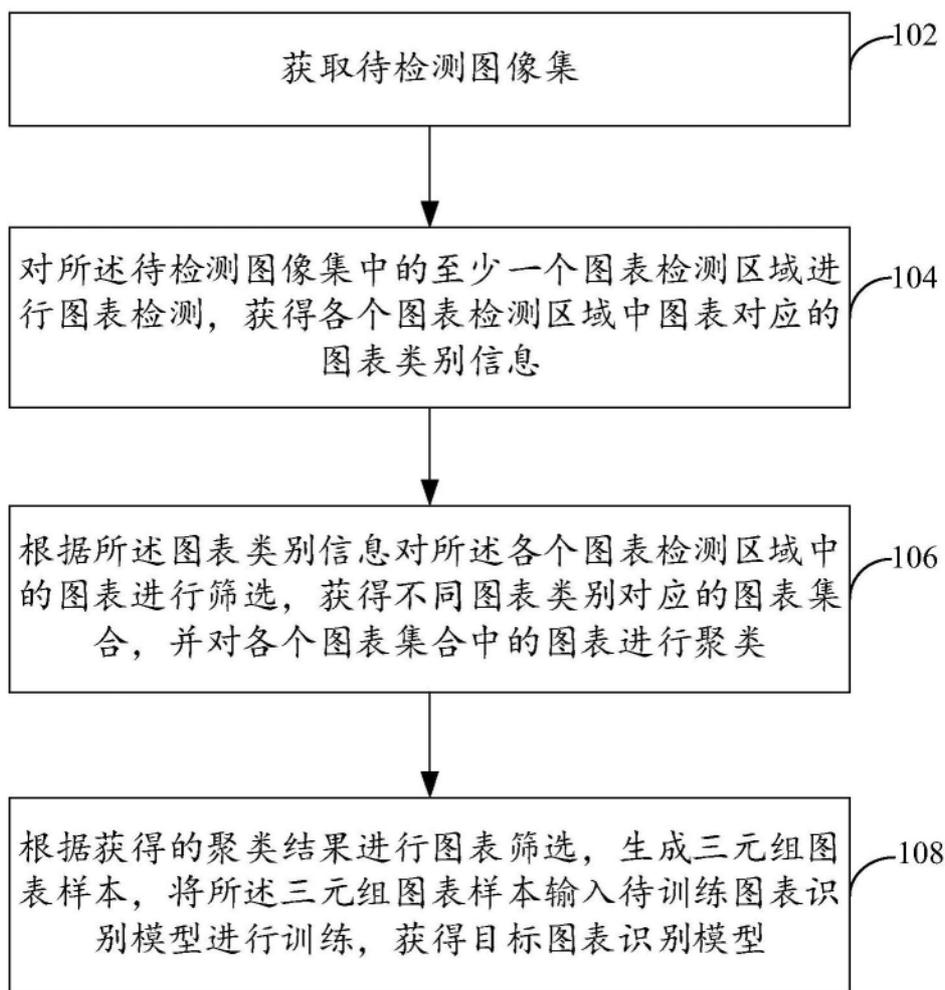


图1

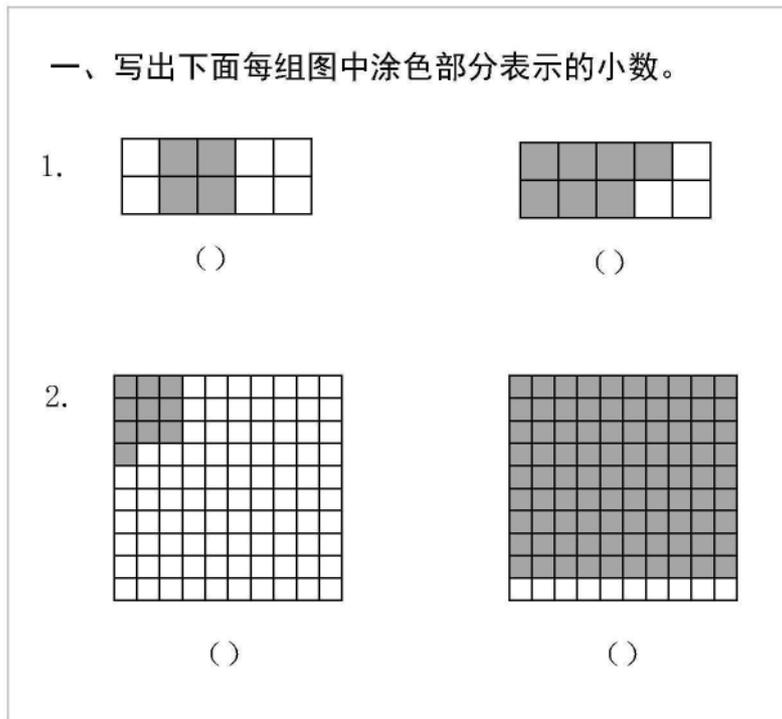


图2a

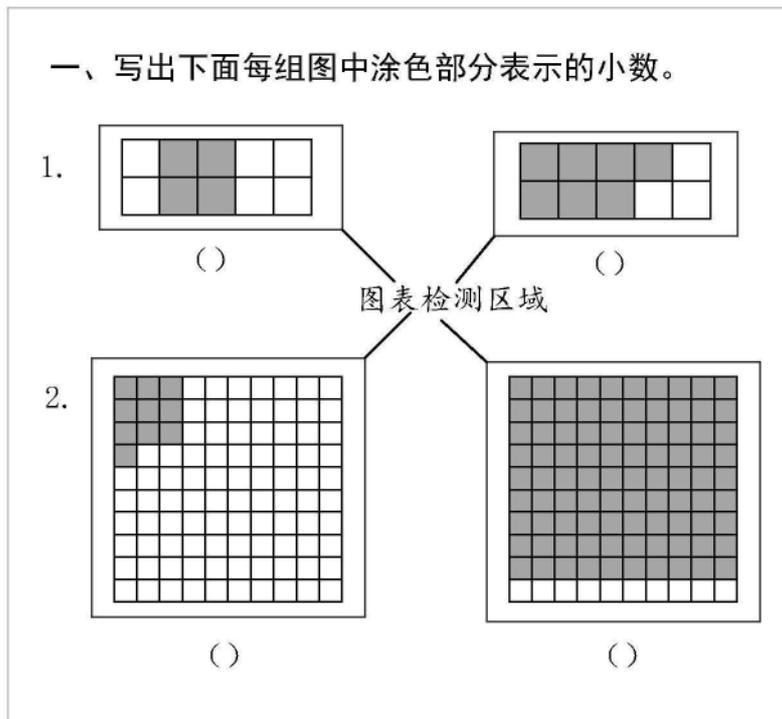


图2b

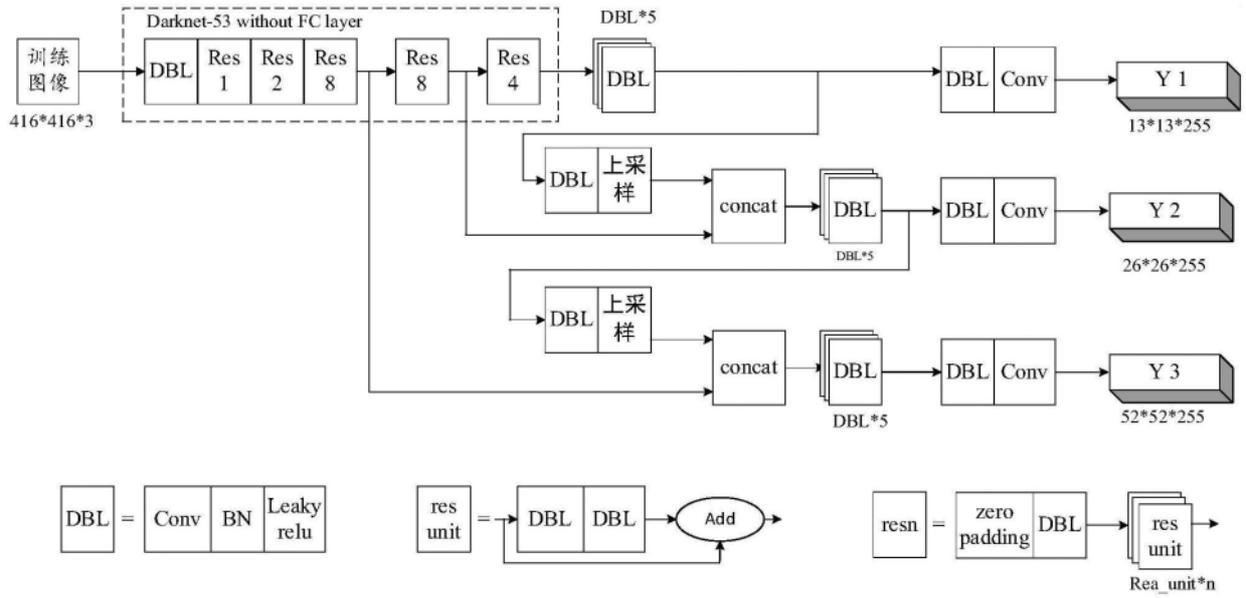


图3

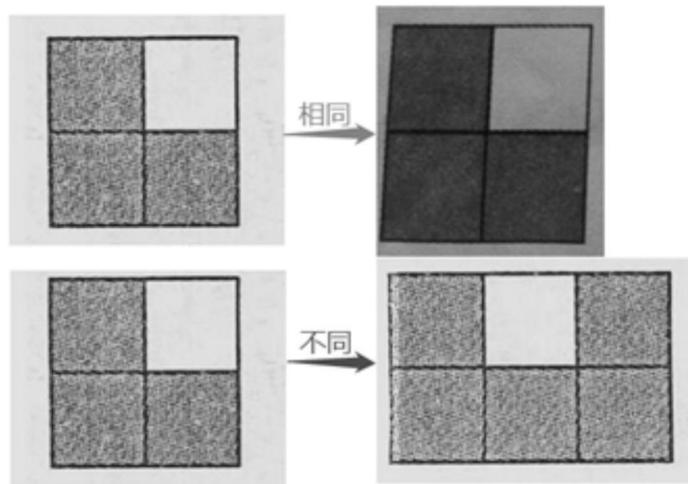


图4

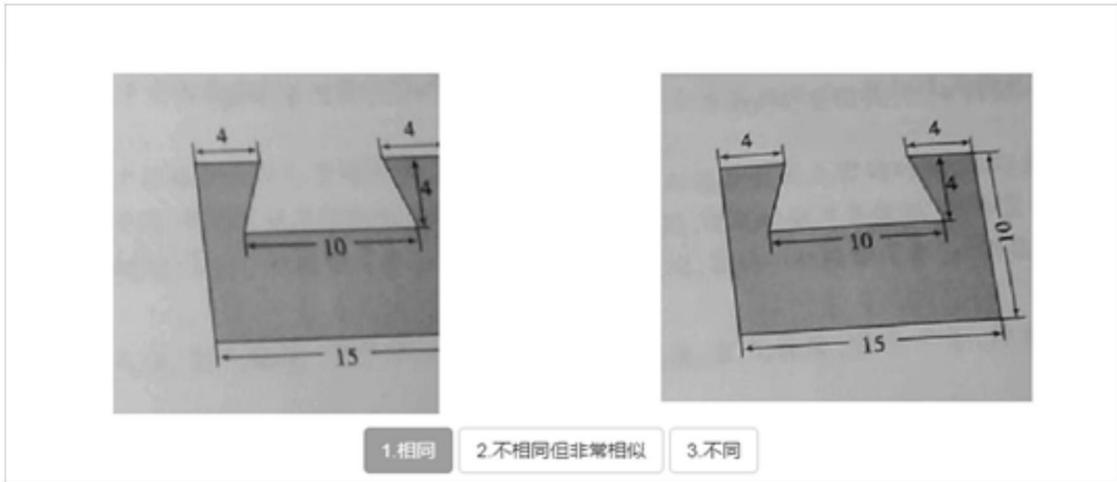


图5a



图5b

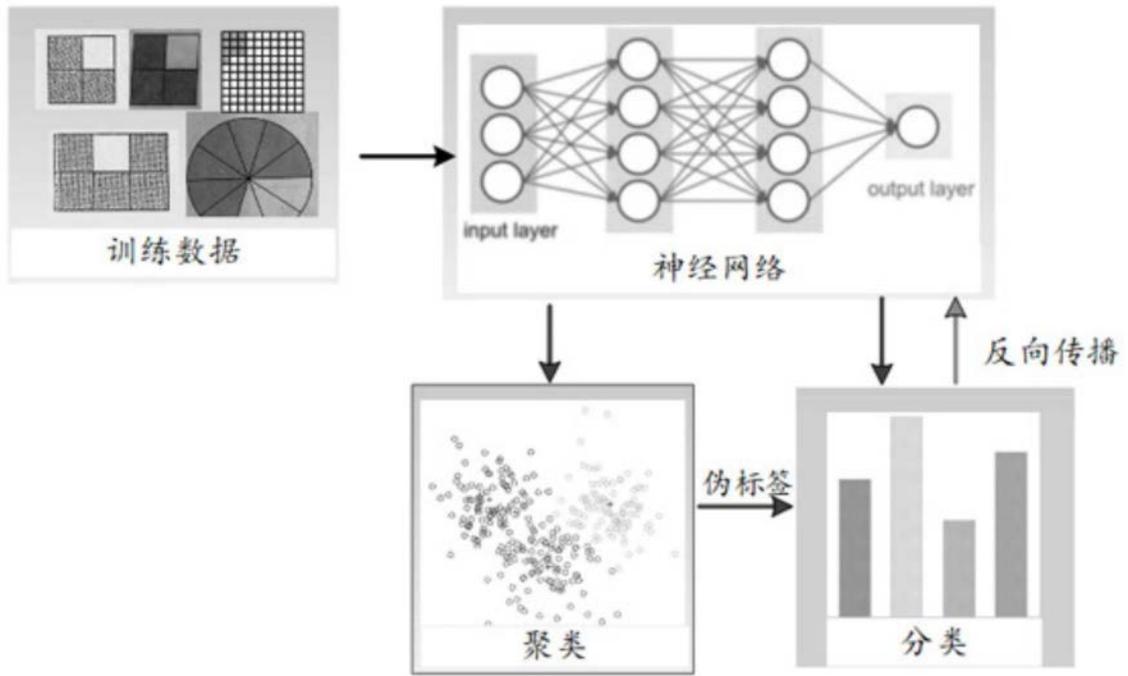


图6

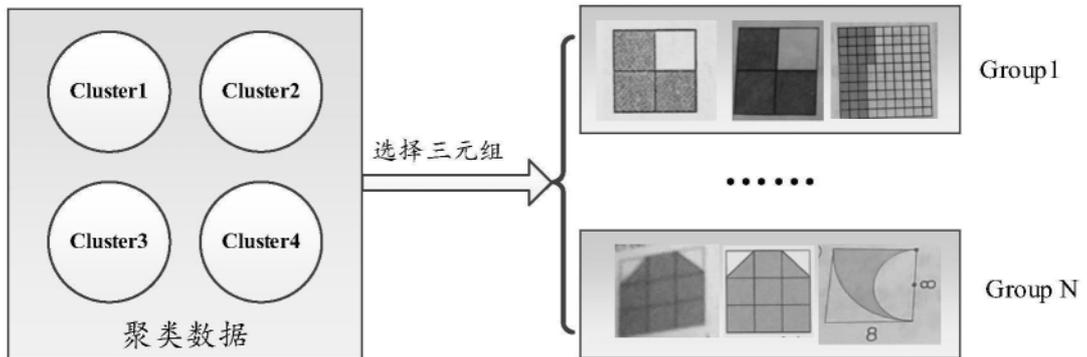


图7

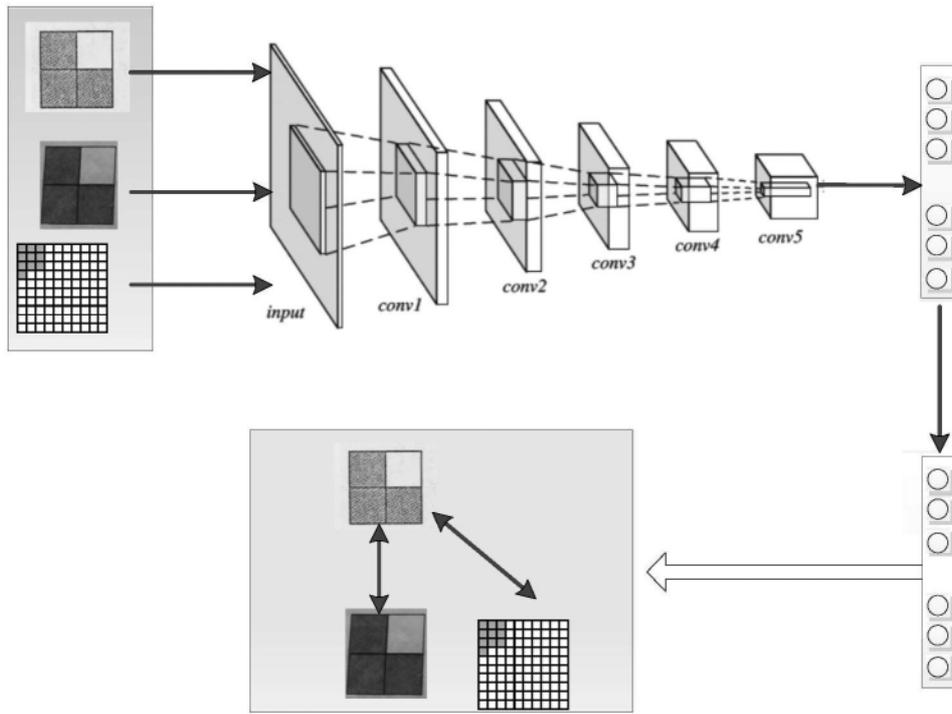


图8

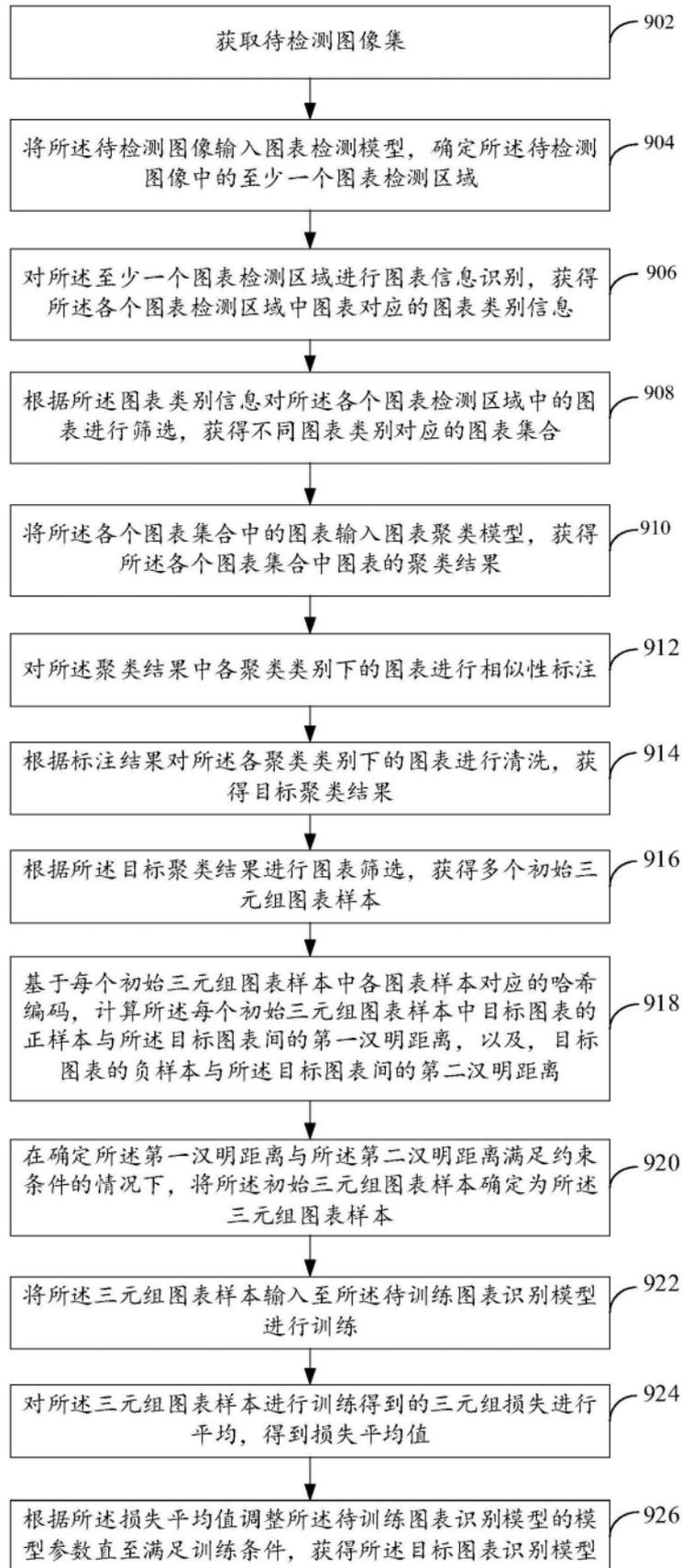


图9

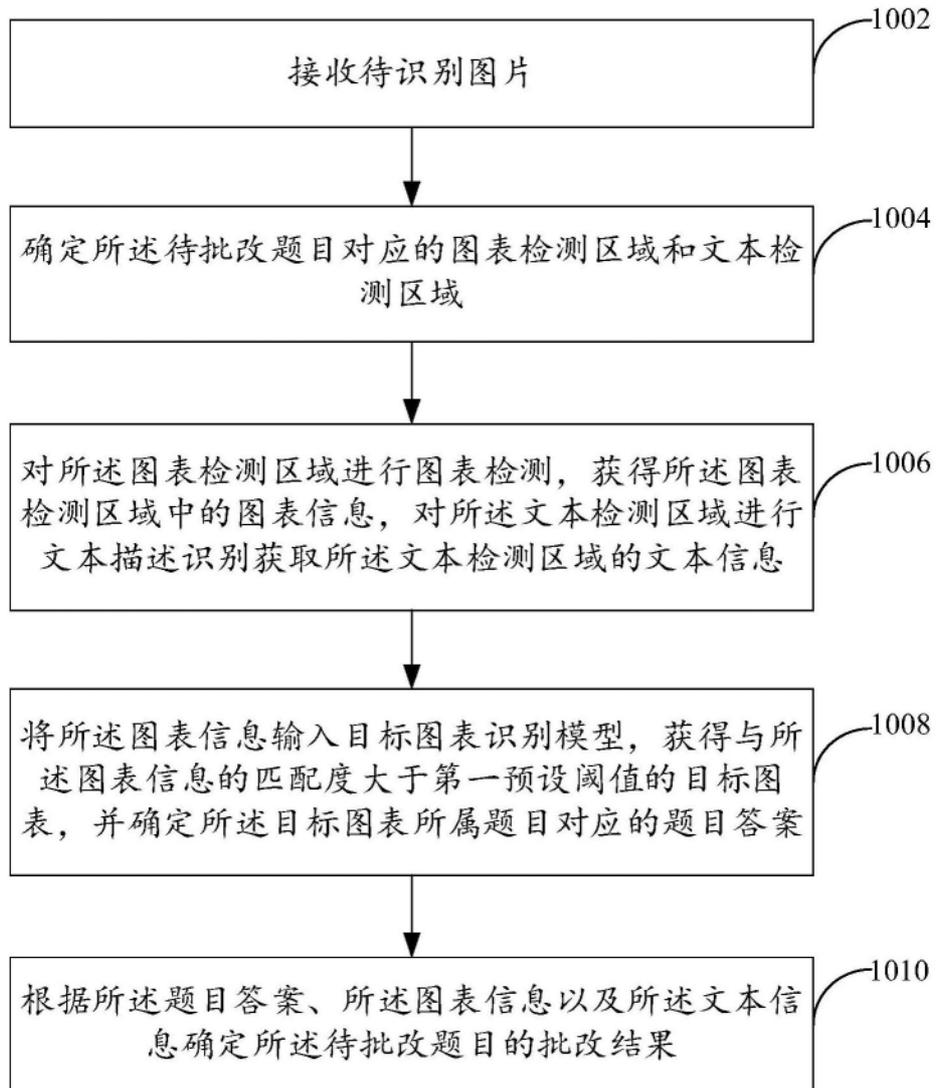


图10

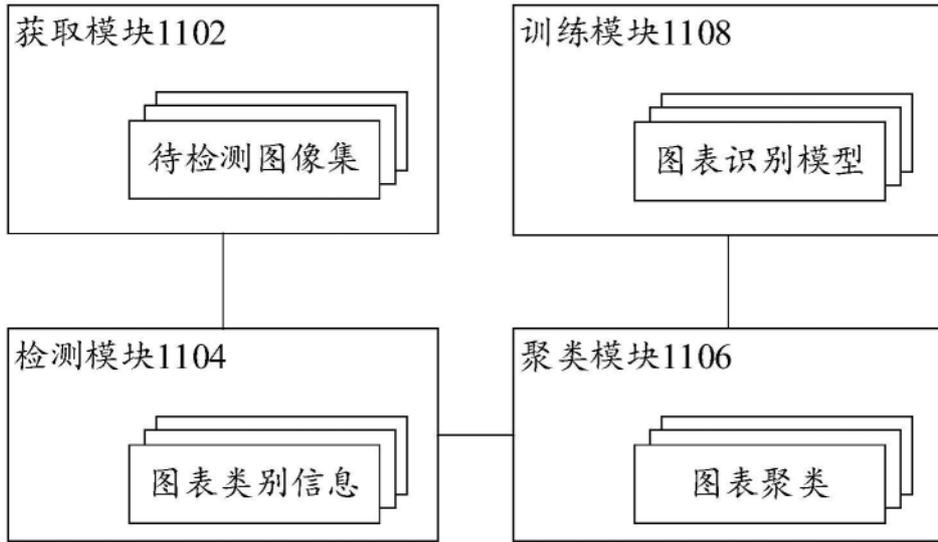


图11

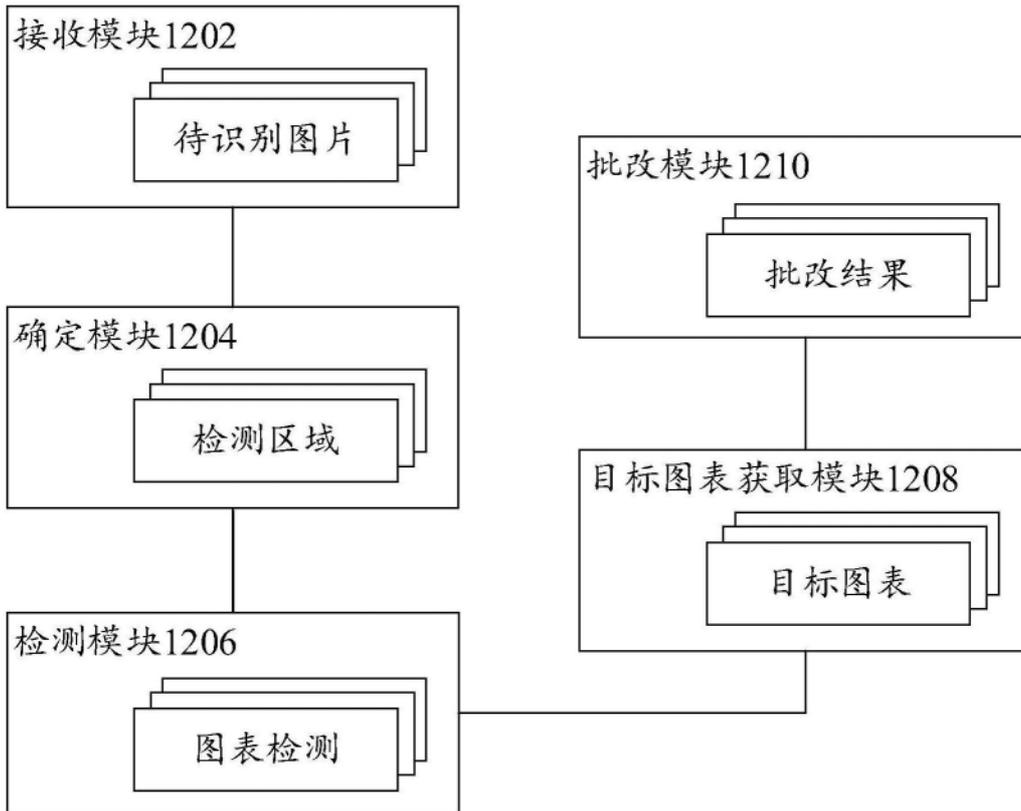


图12

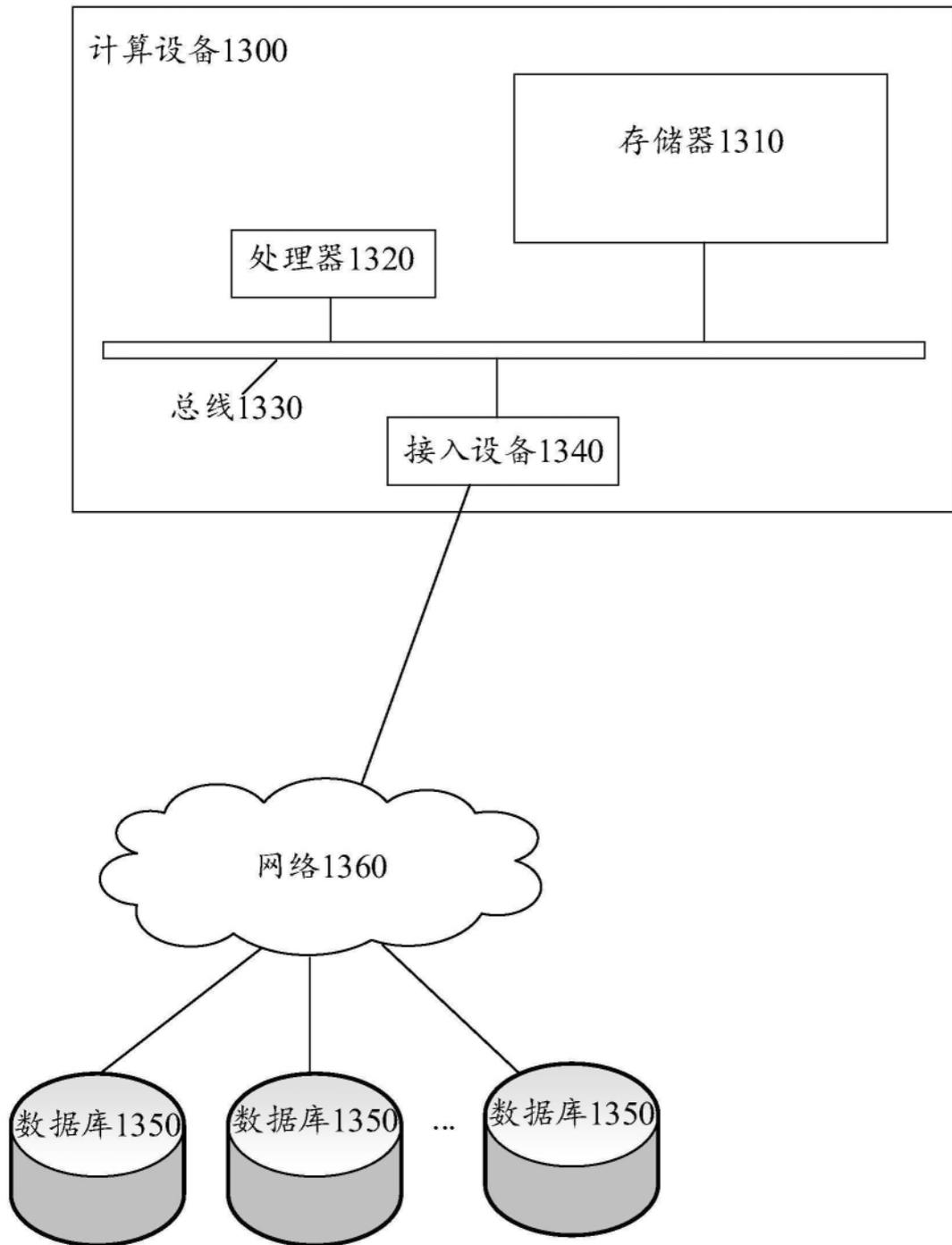


图13