



(12) 发明专利申请

(10) 申请公布号 CN 119941650 A

(43) 申请公布日 2025. 05. 06

(21) 申请号 202411978682.3

G06V 10/80 (2022.01)

(22) 申请日 2024.12.31

G06V 10/25 (2022.01)

(71) 申请人 东莞理工学院

地址 523808 广东省东莞市松山湖科技产业园区大学路1号

(72) 发明人 丁焯 黎鸿铭 陆丽 武雅萍
高子彬

(74) 专利代理机构 广州三环专利商标代理有限公司 44202

专利代理师 杨攀

(51) Int. Cl.

G06T 7/00 (2017.01)

G06T 7/10 (2017.01)

G06V 20/62 (2022.01)

G06V 10/40 (2022.01)

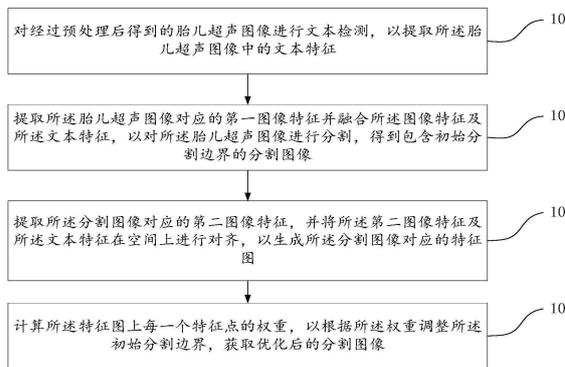
权利要求书2页 说明书11页 附图5页

(54) 发明名称

一种用于胎儿超声图像的多模态分割方法、系统及设备

(57) 摘要

本发明公开了一种用于胎儿超声图像的多模态分割方法、系统及设备,包括:对经过预处理后得到的胎儿超声图像进行文本检测,以提取所述胎儿超声图像中的文本特征;提取所述胎儿超声图像对应的第一图像特征并融合所述图像特征及所述文本特征,以对所述胎儿超声图像进行分割,得到包含初始分割边界的分割图像;提取所述分割图像对应的第二图像特征,并将所述第二图像特征及所述文本特征在空间上进行对齐,以生成所述分割图像对应的特征图;计算所述特征图上每一个特征点的权重,以根据所述权重调整所述初始分割边界,获取优化后的分割图像,以提高图像分割的精准度。



1. 一种用于胎儿超声图像的多模态分割方法,其特征在于,包括:

对经过预处理后得到的胎儿超声图像进行文本检测,以提取所述胎儿超声图像中的文本特征;

提取所述胎儿超声图像对应的第一图像特征并融合所述图像特征及所述文本特征,以对所述胎儿超声图像进行分割,得到包含初始分割边界的分割图像;

提取所述分割图像对应的第二图像特征,并将所述第二图像特征及所述文本特征在空间上进行对齐,以生成所述分割图像对应的特征图;

计算所述特征图上每一个特征点的权重,以根据所述权重调整所述初始分割边界,获取优化后的分割图像。

2. 根据权利要求1所述的一种用于胎儿超声图像的多模态分割方法,其特征在于,所述对经过预处理后得到的胎儿超声图像进行文本检测,以提取所述胎儿超声图像中的文本特征,包括:

通过预设的文本检测算法对所述胎儿超声图像进行文本定位,得到所述胎儿超声图像中包含文本的若干个文本区域及每一个所述文本区域内的若干个字符;

根据所述文本区域及所述字符,通过预训练的语言识别模型提取每一个所述文本区域对应的文本特征,以根据所述文本特征输出每一个所述文本区域对应的结构化文本特征;其中,所述结构化文本特征包括每一个所述文本区域内的文本内容、所述文本区域在图像中的位置、所述文本内容的文本类型。

3. 根据权利要求2所述的一种用于胎儿超声图像的多模态分割方法,其特征在于,所述提取所述胎儿超声图像对应的第一图像特征并融合所述图像特征及所述文本特征,包括:

将所述胎儿超声图像输入至训练好的图像分割模型中,以提取所述胎儿超声图像对应的图像特征;

通过预设与所述图像分割模型中的自注意力机制融合所述图像特征及所述结构化文本特征,以获取所述胎儿超声图像对应的联合特征图。

4. 根据权利要求3所述的一种用于胎儿超声图像的多模态分割方法,其特征在于,所述对所述胎儿超声图像进行分割,得到包含初始分割边界的分割图像,包括:

基于预设与所述图像分割模型中的收缩路径对所述联合特征图进行处理,获得所述联合特征图在不同分辨率上对应的若干个特征图;

计算所述特征图中每一个位置的注意力权重,以根据所述权重对所述特征图进行加权,生成每一个分辨率对应的加权特征图;

通过跳跃连接的方式将不同分辨率对应的加权特征图进行拼接,以根据拼接后重塑的特征图生成所述胎儿超声图像对应的分割掩膜;

根据所述分割掩膜对所述胎儿超声图像进行分割,生成包含初始分割边界的分割图像。

5. 根据权利要求2-3任一项所述的一种用于胎儿超声图像的多模态分割方法,其特征在于,所述提取所述分割图像对应的第二图像特征,包括:

根据结构化文本特征获取所述胎儿图像中的若干个关键文本内容及每一个所述关键文本内容在所述分割图像中的关键文本区域;

从所述分割图像中匹配所述关键文本区域对应的关键图像区域;

将所述关键图像区域输入至训练好的图像分割模型,以提取所述关键图像区域对应的第二图像特征。

6. 根据权利要求5所述的一种用于胎儿超声图像的多模态分割方法,其特征在于,所述将所述第二图像特征及所述文本特征在空间上进行对齐,以生成所述分割图像对应的特征图,包括:

提取所述关键文本区域对应的关键文本特征,并将所述关键文本特征及所述第二图像特征映射到相同的特征空间,以得到映射后的映射文本特征及映射图像特征;

通过几何变换对齐所述映射文本特征及所述映射图像特征,以生成所述关键图像区域对应的特征图。

7. 根据权利要求6所述的一种用于胎儿超声图像的多模态分割方法,其特征在于,所述计算所述特征图上每一个特征点的权重,以根据所述权重调整所述初始分割边界,包括:

基于注意力机制计算所述关键图像区域对应的特征图上每个特征点的注意力权重;

根据所述注意力权重对所述特征图进行加权,并根据加权后的加权特征图对所述关键图像区域的初始分割边界进行调整。

8. 一种用于胎儿超声图像的多模态分割系统,其特征在于,包括文本检测模块、图像分割模块、多模态融合模块及分割优化模块;

所述文本检测模块用于对经过预处理后得到的胎儿超声图像进行文本检测,以提取所述胎儿超声图像中的文本特征;

所述图像分割模块用于提取所述胎儿超声图像对应的第一图像特征并融合所述图像特征及所述文本特征,以对所述胎儿超声图像进行分割,得到包含初始分割边界的分割图像;

所述多模态融合模块用于提取所述分割图像对应的第二图像特征,并将所述第二图像特征及所述文本特征在空间上进行对齐,以生成所述分割图像对应的特征图;

所述分割优化模块用于计算所述特征图上每一个特征点的权重,以根据所述权重调整所述初始分割边界,获取优化后的分割图像。

9. 根据权利要求8所述的一种用于胎儿超声图像的多模态分割系统,其特征在于,所述文本检测模块包括文本识别单元及特征提取单元;

所述文本识别单元用于通过预设的文本检测算法对所述胎儿超声图像进行文本定位,得到所述胎儿超声图像中包含文本的若干个文本区域及每一个所述文本区域内的若干个字符;

所述特征提取单元用于根据所述文本区域及所述字符,通过预训练的语言识别模型提取每一个所述文本区域对应的文本特征,以根据所述文本特征输出每一个所述文本区域对应的结构化文本特征;其中,所述结构化文本特征包括每一个所述文本区域内的文本内容、所述文本区域在图像中的位置、所述文本内容的文本类型。

10. 一种电子设备,其特征在于,包括:一个或多个处理器;

存储装置,用于存储一个或多个程序,当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1至7中任一项所述的一种用于胎儿超声图像的多模态分割方法。

一种用于胎儿超声图像的多模态分割方法、系统及设备

技术领域

[0001] 本发明涉及图像处理技术领域,具体地,涉及一种用于胎儿超声图像的多模态分割方法、系统及设备。

背景技术

[0002] 现代医学科技中,胎儿超声图像在胎儿健康监测中扮演着重要角色,医生通过观察超声图像可对胎儿的形态结构、器官发育情况以及胎儿的活动状态进行检测。

[0003] 在利用超声图像监测胎儿不同的器官或身体结构的健康状况时,现有技术通常依赖于专业人员手动分割和解读超声图像或通过采用自动图像分析技术如基于阈值的分割方法、边缘检测算法、区域生长算法等对超声图像进行分割,以得到胎儿的各个器官结构在超声图像中的区域。

[0004] 在采用上述超声图像分割方法分割胎儿超声图像的过程中,因为超声图像受到多种因素的影响,比如设备差异、操作者技术水平和患者体型等因素导致超声图像的变化性较大,而上述方法因其依赖特定特征和预设参数的操作方式进行图像分割,因此在处理胎儿超声图像时,这些方法难以准确捕捉和分割细微的、高度变化的结构特征,导致最终的分割质量不理想。

发明内容

[0005] 为了解决上述技术问题,本发明公开了一种用于胎儿超声图像的多模态分割方法、系统及设备,用于提高图像分割的精准度。

[0006] 为了实现上述目的,第一方面,本发明公开了一种用于胎儿超声图像的多模态分割方法,包括:

[0007] 对经过预处理后得到的胎儿超声图像进行文本检测,以提取所述胎儿超声图像中的文本信息;

[0008] 提取所述胎儿超声图像对应的第一图像特征并融合所述图像特征及所述文本特征,以对所述胎儿超声图像进行分割,得到包含初始分割边界的分割图像;

[0009] 提取所述分割图像对应的第二图像特征,并将所述第二图像特征及所述文本特征在空间上进行对齐,以生成所述分割图像对应的特征图;

[0010] 计算所述特征图上每一个特征点的权重,以根据所述权重调整所述初始分割边界,获取优化后的分割图像。

[0011] 本发明公开的一种用于胎儿超声图像的多模态分割方法,在对胎儿超声图像进行分割时,首先对所述胎儿超声图像进行文本识别,以提取所述胎儿超声图像中的文本信息,以利用所述文本信息增强后续进行图像分割的精准度。在进行图像的初步分割时,根据所述提取的文本信息及所述胎儿超声图像的第一图像特征进行图像分割,以提高分割的效率及精准度。接着在得到包含初始分割边界的分割图像后,提取所述分割图像的图像特征,并通过对齐所述图像特征及所述文本特征对所述分割图像的分割边界进行优化,进一步提高

图像分割的精准度。

[0012] 作为优选例子,所述对经过预处理后得到的胎儿超声图像进行文本检测,以提取所述胎儿超声图像中的文本特征,包括:

[0013] 通过预设的文本检测算法对所述胎儿超声图像进行文本定位,得到所述胎儿超声图像中包含文本的若干个文本区域及每一个所述文本区域内的若干个字符;

[0014] 根据所述文本区域及所述字符,通过预训练的语言识别模型提取每一个所述文本区域对应的文本特征,以根据所述文本特征输出每一个所述文本区域对应的结构化文本特征;其中,所述结构化文本特征包括每一个所述文本区域内的文本内容、所述文本区域在图像中的位置、所述文本内容的文本类型。

[0015] 上述方案中通过对胎儿超声图像中的文本内容进行检测,为后续进行图像分割时提供进行分割的文本语义信息,进而提高图像分割的精准度。

[0016] 作为优选例子,所述提取所述胎儿超声图像对应的第一图像特征并融合所述图像特征及所述文本特征,包括:

[0017] 将所述胎儿超声图像输入至训练好的图像分割模型中,以提取所述胎儿超声图像对应的图像特征;

[0018] 通过预设 in 所述图像分割模型中的自注意力机制融合所述图像特征及所述结构化文本特征,以获取所述胎儿超声图像对应的联合特征图。

[0019] 上述方案中利用模型进行图像分割,可自适应地根据图像本身的特征进行图像分割,其能捕捉到图像中细微的图像特征,进而提高图像分割的效果,进一步地,在利用模型进行图像分割时,结合文本特征及图像特征将超声图像中的文本信息与图像内容进行一一对应,可提高图像分割的精准度。

[0020] 作为优选例子,所述对所述胎儿超声图像进行分割,得到包含初始分割边界的分割图像,包括:

[0021] 基于预设 in 所述图像分割模型中的收缩路径对所述联合特征图进行处理,获得所述联合特征图在不同分辨率上对应的若干个特征图;

[0022] 计算所述特征图中每一个位置的注意力权重,以根据所述权重对所述特征图进行加权,生成每一个分辨率对应的加权特征图;

[0023] 通过跳跃连接的方式将不同分辨率对应的加权特征图进行拼接,以根据拼接后重塑的特征图生成所述胎儿超声图像对应的分割掩膜;

[0024] 根据所述分割掩膜对所述胎儿超声图像进行分割,生成包含初始分割边界的分割图像。

[0025] 上述方案中,模型利用融合后的特征对图像进行分割时,通过所述收缩途径及所述跳跃连接的方式提取融合后的特征在不同层次中的特征表现,提高特征提取的精准度,进而利用提取的精细的特征对图像进行分割,保证所述分割图像的精准度。

[0026] 作为优选例子,所述提取所述分割图像对应的第二图像特征,包括:

[0027] 根据结构化文本特征获取所述胎儿图像中的若干个关键文本内容及每一个所述关键文本内容在所述分割图像中的关键文本区域;

[0028] 从所述分割图像中匹配所述关键文本区域对应的关键图像区域;

[0029] 将所述关键图像区域输入至训练好的图像分割模型,以提取所述关键图像区域对

应的第二图像特征。

[0030] 上述方案中首先从文本内容中提取出关键文本区域及匹配所述关键文本区域所在的关键图像区域,以在后续通过对应所述关键文本区域和所述关键图像区域对所述分割图像中的关键区域的分割边界进行优化,提高图像分割的精准度。

[0031] 作为优选例子,所述将所述第二图像特征及所述文本特征在空间上进行对齐,以生成所述分割图像对应的特征图,包括:

[0032] 提取所述关键文本区域对应的关键文本特征,并将所述关键文本特征及所述第二图像特征映射到相同的特征空间,以得到映射后的映射文本特征及映射图像特征;

[0033] 通过几何变换对齐所述映射文本特征及所述映射图像特征,以生成所述关键图像区域对应的特征图。

[0034] 上述方案中,通过将对应的文本内容及关键图像的特征进行映射和对齐,以根据所述文本内容对分割后的图像中的关键区域的边界做进一步优化,提高图像分割的精准度。

[0035] 作为优选例子,所述计算所述特征图上每一个特征点的权重,以根据所述权重调整所述初始分割边界,包括:

[0036] 基于注意力机制计算所述关键图像区域对应的特征图上每个特征点的注意力权重;

[0037] 根据所述注意力权重对所述特征图进行加权,并根据加权后的加权特征图对所述关键图像区域的初始分割边界进行调整。

[0038] 上述方案中,引入注意力机制,计算特征对齐后特征图上每一点的注意力权重,突出重要特征,并将权重应用于原始特征,生成加权后的特征表示,不仅提高了图像分割的精度,还增强了对关键区域的关注度,提高图像分割的效果。

[0039] 第二方面,本发明公开了一种用于胎儿超声图像的多模态分割系统,包括文本检测模块、图像分割模块、多模态融合模块及分割优化模块;

[0040] 所述文本检测模块用于对经过预处理后得到的胎儿超声图像进行文本检测,以提取所述胎儿超声图像中的文本特征;

[0041] 所述图像分割模块用于提取所述胎儿超声图像对应的第一图像特征并融合所述图像特征及所述文本特征,以对所述胎儿超声图像进行分割,得到包含初始分割边界的分割图像;

[0042] 所述多模态融合模块用于提取所述分割图像对应的第二图像特征,并将所述第二图像特征及所述文本特征在空间上进行对齐,以生成所述分割图像对应的特征图;

[0043] 所述分割优化模块用于计算所述特征图上每一个特征点的权重,以根据所述权重调整所述初始分割边界,获取优化后的分割图像。

[0044] 本发明公开的一种用于胎儿超声图像的多模态分割系统,在对胎儿超声图像进行分割时,首先对所述胎儿超声图像进行文本识别,以提取所述胎儿超声图像中的文本信息,以利用所述文本信息增强后续进行图像分割的精准度。在进行图像的初步分割时,根据所述提取的文本信息及所述胎儿超声图像的第一图像特征进行图像分割,以提高分割的效率及精准度。接着在得到包含初始分割边界的分割图像后,提取所述分割图像的图像特征,并通过对齐所述图像特征及所述文本特征对所述分割图像的分割边界进行优化,进一步提高

图像分割的精准度。

[0045] 作为优选例子,所述文本检测模块包括文本识别单元及特征提取单元;

[0046] 所述文本识别单元用于通过预设的文本检测算法对所述胎儿超声图像进行文本定位,得到所述胎儿超声图像中包含文本的若干个文本区域及每一个所述文本区域内的若干个字符;

[0047] 所述特征提取单元用于根据所述文本区域及所述字符,通过预训练的语言识别模型提取每一个所述文本区域对应的文本特征,以根据所述文本特征输出每一个所述文本区域对应的结构化文本特征;其中,所述结构化文本特征包括每一个所述文本区域内的文本内容、所述文本区域在图像中的位置、所述文本内容的文本类型。

[0048] 上述方案中通过对胎儿超声图像中的文本内容进行检测,为后续进行图像分割时提供进行分割的文本语义信息,进而提高图像分割的精准度。

[0049] 第三方面,本发明公开了一种电子设备,包括:一个或多个处理器;存储装置,用于存储一个或多个程序,当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如第一方面所述的一种用于胎儿超声图像的多模态分割方法。

附图说明

[0050] 图1:为本发明实施例公开的一种用于胎儿超声图像的多模态分割方法的流程示意图;

[0051] 图2:为本发明实施例公开的一种用于胎儿超声图像的多模态分割系统的结构示意图;

[0052] 图3:为本发明又一实施例公开的一种用于胎儿超声图像的多模态分割方法的流程示意图;

[0053] 图4:为本发明又一实施例公开的一种模型训练流程示意图;

[0054] 图5:为本发明又一实施例公开的一种文本识别流程示意图;

[0055] 图6:为本发明又一实施例公开的一种图像分割流程示意图;

[0056] 图7:为本发明又一实施例公开的一种图像分割优化流程示意图。

具体实施方式

[0057] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0058] 实施例一

[0059] 参照图1,为本发明实施例公开的一种用于胎儿超声图像的多模态分割方法,用于提高图像分割的精准度。具体地,所述多模态分割方法的具体实施流程请参照图1,主要包括步骤101至步骤104,所述步骤为:

[0060] 步骤101:对经过预处理后得到的胎儿超声图像进行文本检测,以提取所述胎儿超声图像中的文本特征。

[0061] 步骤102:提取所述胎儿超声图像对应的第一图像特征并融合所述图像特征及所

述文本特征,以对所述胎儿超声图像进行分割,得到包含初始分割边界的分割图像。

[0062] 步骤103:提取所述分割图像对应的第二图像特征,并将所述第二图像特征及所述文本特征在空间上进行对齐,以生成所述分割图像对应的特征图。

[0063] 步骤104:计算所述特征图上每一个特征点的权重,以根据所述权重调整所述初始分割边界,获取优化后的分割图像。

[0064] 在本公开实施例中,在经过超声扫描获得胎儿超声图像后,为根据所述胎儿超声图像获取胎儿的器官、身体结构等的健康状态,需要对所述胎儿超声图像进行图像分割,以生成所述胎儿超声图像对应的分割边界,进而从所述分割边界内包围的每一干个单独区域内对应的结构图像进行健康检测。其中,基于现有的用于生成超声图像的设备在生成胎儿超声图像时,会接收用户的输入或自动地根据扫描的进程在所述超声图像中的每一区域进行标注,以方便根据所述标注区分图像的不同区域。对此,在进行图像分割,可先对接收的胎儿超声图像进行文本检测,识别所述超声图像中的文本信息,以根据所述文本信息进行图像分割,进而提高图像分割的精准度。

[0065] 具体地,在提取所述胎儿超声图像中的文本信息时,步骤101可提取所述图像的文本特征,进而根据所述文本特征识别对应的文本信息。优选地,所述步骤101可通过以下步骤进行文本特征的识别:

[0066] 步骤1011:通过预设的文本检测算法对所述胎儿超声图像进行文本定位,得到所述胎儿超声图像中包含文本的若干个文本区域及每一个所述文本区域内的若干个字符;

[0067] 步骤1012根据所述文本区域及所述字符,通过预训练的语言识别模型提取每一个所述文本区域对应的文本特征,以根据所述文本特征输出每一个所述文本区域对应的结构化文本特征;其中,所述结构化文本特征包括每一个所述文本区域内的文本内容、所述文本区域在图像中的位置、所述文本内容的文本类型。

[0068] 在本公开实施例中,上述步骤通过对胎儿超声图像中的文本内容进行检测,为后续进行图像分割时提供进行分割的文本语义信息,进而提高图像分割的精准度。

[0069] 进一步地,在提取对应的文本特征后,为自适应地对不同的超声图像进行图像分割,可基于深度学习构建模型,以自适应学习不同图像中细微的结构特征,进而利用训练好的模型进行图像分割。具体地,在利用模型进行图像分割的过程中,步骤102可通过以下步骤融合图像的特征及文本特征进行图像分割,优选地:

[0070] 步骤1021:将所述胎儿超声图像输入至训练好的图像分割模型中,以提取所述胎儿超声图像对应的图像特征;

[0071] 步骤1022:通过预设在所述图像分割模型中的自注意力机制融合所述图像特征及所述结构化文本特征,以获取所述胎儿超声图像对应的联合特征图。

[0072] 在本公开实施例中,上述步骤利用模型进行图像分割,可自适应地根据图像本身的特征进行图像分割,其能捕捉到图像中细微的图像特征,进而提高图像分割的效果,进一步地,在利用模型进行图像分割时,结合文本特征及图像特征将超声图像中的文本信息与图像内容进行一一对应,可提高图像分割的精准度。

[0073] 具体地,在步骤1022通过融合后的特征进行图像分割时,为保证融合的精准度,则所述步骤1022通过以下步骤进行图像分割,包括:

[0074] 步骤10221:基于预设在所述图像分割模型中的收缩路径对所述联合特征图进行

处理,获得所述联合特征图在不同分辨率上对应的若干个特征图;

[0075] 步骤10222:计算所述特征图中每一个位置的注意力权重,以根据所述权重对所述特征图进行加权,生成每一个分辨率对应的加权特征图;

[0076] 步骤10223:通过跳跃连接的方式将不同分辨率对应的加权特征图进行拼接,以根据拼接后重塑的特征图生成所述胎儿超声图像对应的分割掩膜;

[0077] 步骤10224:根据所述分割掩膜对所述胎儿超声图像进行分割,生成包含初始分割边界的分割图像。

[0078] 在本公开实施例中,上述步骤利用融合后的特征对图像进行分割时,通过所述收缩途径及所述跳跃连接的方式提取融合后的特征在不同层次中的特征表现,提高特征提取的精准度,进而利用提取的精细的特征对图像进行分割,保证所述分割图像的精准度。

[0079] 进一步地,在通过图像分割模型输出包含初始分割边界的分割图像后,为了进一步提高图像分割的精准度,可通过匹配文本内容及分割图像中的分割结果对所述分割图像做进一步优化,具体地,所述步骤103通过提取分割图像对应的图像特征及文本内容的文本特征进行文本与图像的匹配,所述步骤103主要包括:

[0080] 步骤1031:根据结构化文本特征获取所述胎儿图像中的若干个关键文本内容及每一个所述关键文本内容在所述分割图像中的关键文本区域;

[0081] 步骤1032:从所述分割图像中匹配所述关键文本区域对应的关键图像区域;

[0082] 步骤1033:将所述关键图像区域输入至训练好的图像分割模型,以提取所述关键图像区域对应的第二图像特征;

[0083] 步骤1034:提取所述关键文本区域对应的关键文本特征,并将所述关键文本特征及所述第二图像特征映射到相同的特征空间,以得到映射后的映射文本特征及映射图像特征;

[0084] 步骤1035:通过几何变换对齐所述映射文本特征及所述映射图像特征,以生成所述关键图像区域对应的特征图。

[0085] 在本公开实施例中,上述步骤首先从文本内容中提取出关键文本区域及匹配所述关键文本区域所在的关键图像区域,以在后续通过对应所述关键文本区域和所述关键图像区域对所述分割图像中的关键区域的分割边界进行优化,提高图像分割的精准度,进一步地,通过将对应的文本内容及关键图像的特征进行映射和对齐,以根据所述文本内容对分割后的图像中的关键区域的边界做进一步优化,提高图像分割的精准度。

[0086] 进一步地,在对齐所述文本内容与所述图像的特征后,根据对齐后生成的特征图对所述初始分割边界进行优化,具体地,所述步骤104通过以下步骤对所述分割边界进行优化,包括:

[0087] 步骤1041:基于注意力机制计算所述关键图像区域对应的特征图上每个特征点的注意力权重;

[0088] 步骤1042:根据所述注意力权重对所述特征图进行加权,并根据加权后的加权特征图对所述关键图像区域的初始分割边界进行调整。

[0089] 在本公开实施例中,上述步骤引入注意力机制,计算特征对齐后特征图上每一点的注意力权重,突出重要特征,并将权重应用于原始特征,生成加权后的特征表示,不仅提高了图像分割的精度,还增强了对关键区域的关注度,提高图像分割的效果。

[0090] 另一方面,本实施例还公开了一种用于胎儿超声图像的多模态分割系统,所述分割系统的具体结构组成请参照图2,包括文本检测模块201、图像分割模块202、多模态融合模块203及分割优化模块204。

[0091] 所述文本检测模块201用于对经过预处理后得到的胎儿超声图像进行文本检测,以提取所述胎儿超声图像中的文本特征。

[0092] 所述图像分割模块202用于提取所述胎儿超声图像对应的第一图像特征并融合所述图像特征及所述文本特征,以对所述胎儿超声图像进行分割,得到包含初始分割边界的分割图像。

[0093] 所述多模态融合模块203用于提取所述分割图像对应的第二图像特征,并将所述第二图像特征及所述文本特征在空间上进行对齐,以生成所述分割图像对应的特征图。

[0094] 所述分割优化模块204用于计算所述特征图上每一个特征点的权重,以根据所述权重调整所述初始分割边界,获取优化后的分割图像。

[0095] 在本实施例中,所述文本检测模块201包括文本识别单元及特征提取单元。

[0096] 所述文本识别单元用于通过预设的文本检测算法对所述胎儿超声图像进行文本定位,得到所述胎儿超声图像中包含文本的若干个文本区域及每一个所述文本区域内的若干个字符。

[0097] 所述特征提取单元用于根据所述文本区域及所述字符,通过预训练的语言识别模型提取每一个所述文本区域对应的文本特征,以根据所述文本特征输出每一个所述文本区域对应的结构化文本特征;其中,所述结构化文本特征包括每一个所述文本区域内的文本内容、所述文本区域在图像中的位置、所述文本内容的文本类型。

[0098] 根据如图1所示的一种用于胎儿超声图像的多模态分割方法,相应的,本实施例提供了一种电子设备,包括处理器、存储器以及存储在所述存储器中且被配置为由所述处理器执行的计算机程序,所述处理器执行所述计算机程序时实现如图1所示的用于胎儿超声图像的多模态分割方法。其中,所称处理器可以是中央处理单元(Central Processing Unit,CPU),还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等,所述处理器是所述终端的控制中心,利用各种接口和线路连接整个终端的各个部分。所述存储器可用于存储所述计算机程序,所述处理器通过运行或执行存储在所述存储器内的计算机程序,以及调用存储在存储器内的数据,实现所述终端的各种功能。所述存储器可主要包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需的应用程序(比如声音播放功能、图像播放功能等)等;存储数据区可存储根据手机的使用所创建的数据(比如音频数据、电话本等)等。此外,存储器可以包括高速随机存取存储器,还可以包括非易失性存储器,例如硬盘、内存、插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)、至少一个磁盘存储器件、闪存器件,或其他易失性固态存储器件。

[0099] 相比于现有技术,本实施例提供的用于胎儿超声图像的多模态分割方法、系统及设备具有如下有益效果:

[0100] 本实施例提供的一种用于胎儿超声图像的多模态分割方法、系统及设备,在对胎儿超声图像进行分割时,首先对所述胎儿超声图像进行文本识别,以提取所述胎儿超声图像中的文本信息,以利用所述文本信息增强后续进行图像分割的精准度。在进行图像的初步分割时,根据所述提取的文本信息及所述胎儿超声图像的第一图像特征进行图像分割,以提高分割的效率及精准度。接着在得到包含初始分割边界的分割图像后,提取所述分割图像的图像特征,并通过对齐所述图像特征及所述文本特征对所述分割图像的分割边界进行优化,进一步提高图像分割的精准度。

[0101] 实施例二

[0102] 现有技术对胎儿超声图像进行分割,以对胎儿的器官、结构等发育情况进行健康检测时,常通过手动分割或自动图像分析技术等对超声图像的分割。其中,采用手动分割对超声图像进行分割时,不仅受限于图像分割人员的专业知识导致图像分割的效率比较低,且人工分割的精准度不高,而采用自动图像分析技术进行图像分割时,自动图像分析技术依赖特定特征和预设参数,无法使用变化性极强的超声图像,导致图像的分割效果不高。

[0103] 对此,为解决上述图像分割方法导致的技术问题,本实施例公开了一种胎儿超声图像的多模态分割方法,利用深度学习构建图像分割模型,以自适应提取不同超声图像的图像特征进行图像分割,提高分割的效果。同时结合图像中的文本信息为图像分割提供依据及对分割后的图像进行优化,提高图像分割的精准度。具体地,本实施例公开的所述多模态分割方法的具体实施流程请参照图3,主要包括步骤301至步骤304:

[0104] 步骤301:基于预构建的数据训练集对预构建的深度学习模型进行训练,以获得图像分割模型及文本识别模型。

[0105] 具体地,在本实施例某些实施方式中,首先采集超声图像数据,进而构建训练数据集以对构建的深度学习模型进行训练,生成用于文本识别的文本识别模型和图像分割模型。

[0106] 优选地,在本实施例某些实施方式中,所述文本识别模型和图像分割模型的训练过程可参照图4。如图4所示,首先进行超声图像数据的采集,其中,可从保存超声图像的数据库或现有的图像数据库中提取若干张历史胎儿超声图像、每一张所述历史胎儿超声图像对应的包含分割边界的分割图像及每一张所述历史胎儿超声图像对应的文本信息。其中,所述文本信息包括所述历史胎儿超声图像中的文本内容、文本内容在图像中的文本区域、文本类型(如标题、段落、注释等)、特殊标记和符号,以及可能的语义标签。

[0107] 进一步地,在进行模型训练的过程中,为保证训练的效率及精准度,可对每一张所述历史胎儿超声图像进行预处理,所述预处理包括对历史胎儿超声图像进行必要的处理,如图像尺寸统一、对比度调整、去噪处理等,以优化后续模型的输入质量、通过使用先进的图像处理算法和工具(如OpenCV、PIL),以对图像进行转换和增强等。

[0108] 然后将经过预处理后的每一张所述历史胎儿超声图像输入至基于深度学习网络构建的初始文本识别模型中,以获取所述初始文本识别模型输出的初始文本信息。其中,可采用基于深度学习的OCR技术、CTPN光符文字检测算法等专门用于图像文字检测的算法构建所述初始文本识别模型。

[0109] 接着,在进行训练的过程中,获取所述初始文本信息和训练集中对应的文本信息的损失,以根据所述损失调整所述初始文本识别模型的模型参数,直至损失小于预设的损

失阈值后,保留当前的模型参数获得文本识别模型。

[0110] 在进行图像分割模型的训练时,将经过预处理后的每一张所述历史胎儿超声图像及所述文本信息输入至基于深度学习网络构建的初始图像分割模型中,以获取所述初始图像分割模型输出的初始分割图像。其中,为提高胎儿超声图像分割的精准度,可采用U-NET网络、FCN网络、DeepLab v3+网络等专门用于医学图像分割的深度学习网络的参数构建所述初始分割图像。

[0111] 在进行训练的过程中,获取所述初始分割图像与训练集中对应的分割图像的损失,以根据所述损失调整所述初始图像分割模型的模型参数,直至损失小于预设的损失阈值后,保留当前的模型参数获得图像分割模型。

[0112] 步骤302:对待分割的胎儿超声图像进行预处理,并将预处理后的胎儿超声图像输入至所述文本识别模型中,以提取所述胎儿超声图像的结构化文本特征。

[0113] 具体地,在本实施例中,基于模型训练过程中对图像进行了预处理操作,则在本实施例中,当通过训练好的文本识别模型提取胎儿超声图像的结构化文本特征时,需对所述胎儿超声图像进行同样的预处理操作,包括对胎儿超声图像进行必要的处理,包括图像尺寸统一、对比度调整、去噪处理等,以优化后续模型的输入质量。通过使用先进的图像处理算法和工具(如OpenCV、P I L),对图像进行转换和增强。

[0114] 进一步地,将经过预处理的胎儿超声图像输入至所述文本识别模型,以通过所述文本识别模型输出所述胎儿超声图像的结构化文本特征的过程可参照图5。

[0115] 如图5所示,在图像输入所述文本识别模型后,首先通过文本识别模型中的深度卷积神经网络对图像进行处理,识别出包含文本的区域。这些文本区域被进一步细化,分割成多个独立的文本块。同时,利用卷积神经网络对每个文本块中的字符进行逐个分割和识别,确保每个字符都能被准确识别和分类。

[0116] 接着,在所述文本识别模型过程中,因对模型进行训练的数据集中为胎儿超声图像领域的文本数据,这些文本数据使模型能够更好地理解和识别医学术语和文本。因此在对所述图像中的文本进行定位后,所述文本识别模型对这些数据进行处理,提取出文本的特征。

[0117] 进一步地,在提取文本的特征时,所述文本识别模型可通过多步骤的校正过程对提取的特征进行优化。首先,通过拼写检查和语法校正来修正识别过程中可能出现的错误。然后,所述文本识别模型利用上下文信息对识别结果进行进一步优化,确保输出的文本准确无误。最终,文本识别模型将提取的特征转化为结构化文本特征。这些结构化特征包括识别出的文本内容、文本在图像中的位置、文本类型(如标题、段落、注释等)、特殊标记和符号,以及可能的语义标签。

[0118] 其中,所述结构化文本特征为后续的文本分析和信息提取提供了坚实的基础,使得用户能够对提取的文本信息进行进一步的处理和利用。例如,在医学图像的应用场景中,可以根据识别出的文本内容自动生成报告、提取关键信息,甚至辅助医生进行诊断和决策。通过这些详细而准确的处理步骤。

[0119] 步骤303:将所述结构化文本特征和经过预处理的胎儿超声图像输入至所述训练好的图像分割模型中,以输出包含初始分割边界的初始分割图像。

[0120] 具体地,在本实施例中,在进行图像分割时,首先加载所述胎儿超声图像和所述结

构化文本特征至所述图像分割模型,以提取所述胎儿超声图像的图像特征,并将所述图像特征和所述结构化文本特征进行融合,进而根据融合后的特征对图像进行分割,以生成包含初始分割边界的分割图像。

[0121] 优选地,所述图像分割模型融合图像特征和文本特征进行图像分割的过程可参照图6。如图6所示,所述图像分割模型首先使用卷积神经网络对输入的胎儿图像进行多层次的特征提取,以生成胎儿超声图像的图像特征。

[0122] 其次,将所述文本识别模型提取的文本特征与提取的图像特征进行深度融合,所述图像特征与文本特征的融合可采用自注意力或全连接层来整合图像和文本的特征表示,以将文本特征中包括的文本内容、文本位置、文本类型等信息与图像特征相结合,生成胎儿超声图像对应的包括图像特征及文本特征的联合特征,以利用文本信息中的上下文和图像中像素的上下文来增强图像分割的准确性和鲁棒性。例如,如果文本信息提到特定的解剖结构,模型可以特别关注这些区域,从而提高分割的精度。

[0123] 接着,参照图6,将融合后的特征输入所述图像分割模型中的收缩路径,所述收缩路径通过一系列卷积层和池化层逐步提取所述联合特征在不同层次的特征图。其中,所述收缩路径中的每一层卷积操作提取特征,并通过池化层降低特征图的尺寸,以获取不同分辨率的特征表示,进而生成不同分辨率分别对应的特征图。

[0124] 进一步地,将所述不同分辨率的特征图传输到所述图像分割模型中的注意力模块,以通过计算特征图中每个位置的注意力权重,增强对分割任务中重要区域的关注度。其中,所述注意力机制可以是通道注意力或空间注意力,通过加权特征图的不同部分,注意力模块通过计算每个特征点的权重来实现对不同位置特征的加权,通常采用通道注意力或空间注意力机制。通道注意力通过聚合特征图的空间信息来评估每个通道的重要性,而空间注意力则直接评估特征图中每个像素位置的重要性。这些权重随后被应用于原始特征图,生成加权后的特征表示,以突出图像中的关键区域。对于每一张输入的特征图,注意力模块都会输出一个对应的加权特征图,这些加权特征图反映了模型对图像中不同区域的关注度,进而用于提高分割任务的准确性。

[0125] 进一步地,对每一个加权后的加权特征图进行上采样操作和特征拼接,使特征图恢复到原始图像的尺寸,同时保留关键的上下文信息,以将收缩路径输出的多张特征图通过特征重塑过程被整合。在特征重塑中,这些多张特征图不是简单地重塑为一张,而是通过上采样和特征拼接的方式,使特征图恢复到原始图像的尺寸,同时保留关键的上下文信息。具体地,这个过程涉及将收缩路径的低分辨率特征图逐步上采样,并通过跳跃连接(Skip Connection)与来自前一层的高分辨率特征图进行拼接,以丰富特征表示并保留细节。跳跃连接特征是通过在U-Net模型的编码器和解码器之间直接连接相同分辨率的特征图来获取的,这样做可以在解码器中直接利用编码器中的详细信息,从而提高分割的准确性和鲁棒性。

[0126] 最后,将特征重塑后的特征图输入至图像分割模型中的扩张路径,以通过逐步上采样和卷积操作,将重塑后的特征图恢复到原始图像的尺寸。每次上采样后,结合来自收缩路径的跳跃连接特征,确保在重建过程中保留更多的细节和边缘信息,最终生成高质量的分割掩膜。

[0127] 最终的分割结果通过扩张路径重建的掩膜,能够准确定位和标识图像中的关键结

构和特征。这些分割结果不仅包括目标区域的精确定位和分割,还能清晰地标识出图像中的关键结构和特征,为后续的分析 and 处理提供基础,从而实现了对医学图像的高效、准确分割。

[0128] 步骤304:提取所述初始分割图像对应的图像特征,以通过对齐所述图像特征和所述文本特征对所述初始分割边界进行优化,以输出优化后的分割图像。

[0129] 具体地,在本实施例中,根据文本特征获取所述胎儿图像中的若干个关键文本内容及每一个所述关键文本内容在所述分割图像中的关键文本区域;从所述分割图像中匹配所述关键文本区域对应的关键图像区域。通过对齐所述关键文本区域对应的文本特征和所述关键图像区域对应的图像特征对所述初始分割边界进行优化。

[0130] 优选地,分割图像优化的过程可参照图7。如图7所示,首先根据文本特征提取关键文本区域及从分割后的超声图像中提取所述关键文本区域对应的关键图像区域。接着,利用文本识别模型和图像分割模型提取关键文本区域的文本特征和关键图像区域的图像特征,并利用卷积神经网络(CNN)和全连接层将这些特征映射到相同的特征空间。特征对齐通过几何变换确保文本特征与图像特征在空间上精确对应。

[0131] 接着,引入注意力机制,计算每个特征点的注意力权重,突出重要特征,并通过点积注意力或加性注意力方法将权重应用于原始特征,生成加权后的特征表示。最终根据所述加权后的特征表示对所述分割图像进行优化。

[0132] 进一步地,在进行图像优化后,还可根据构建的用户反馈机制为用户提供交互式的界面,以使用户可以评估和调整分割结果。确保最终输出的准确性和可用性。其中,用户反馈机制主要是使用户通过界面可以直观地查看分割结果,并进行必要的调整,如标记错误区域或调整分割边界。最终的分割结果和文本信息将被展示和输出,可以直接用于临床诊断或作为医疗记录的一部分。此外,还可进行结果验证,即通过与专家的标注结果进行比较,确保分割结果的准确性和一致性。

[0133] 本实施例公开的一种用于胎儿超声图像的多模态分割方法,采用了多模态融合技术,将识别的文本信息与图像分割能力相结合,能够利用文本信息中蕴含的丰富上下文,指导图像分割过程,从而提高分割的准确性和鲁棒性。与现有技术相比,这种方法不仅提升了分割质量,还能够处理更加复杂的图像场景,尤其是在胎儿超声图像中,能够更准确地识别和定位胎儿的器官和结构。进一步地,利用深度学习模型的强大特征提取和学习能力,显著提高了图像分割的速度和准确度。与依赖于手工特征提取的传统图像处理方法相比,深度学习模型能够自动学习图像中的复杂模式,使得在处理大量数据时,能够快速提供高质量的分割结果,确保了在面对大规模数据集时能快速响应和高准确度输出。最后在结果输出时,充分考虑了用户体验,提供了反馈机制,降低了对专业技术人员的依赖,使得更多的医疗工作者能够直接利用本发明进行有效的医疗影像分析,提高了工作效率,同时也降低了操作的复杂性和出错率。

[0134] 以上所述的具体实施例,对本发明的目的、技术方案和有益效果进行了进一步的详细说明,应当理解,以上所述仅为本发明的具体实施例而已,并不用于限定本发明的保护范围。特别指出,对于本领域技术人员来说,凡在本发明的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

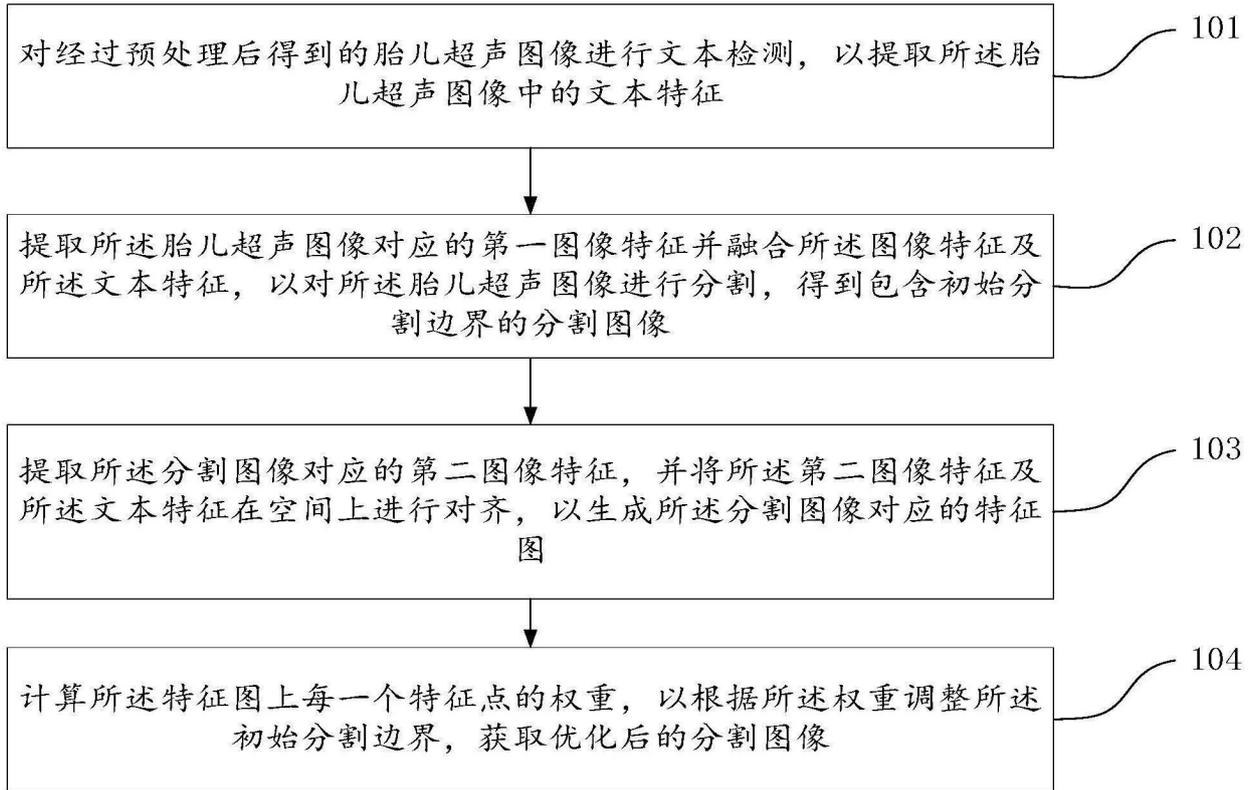


图1

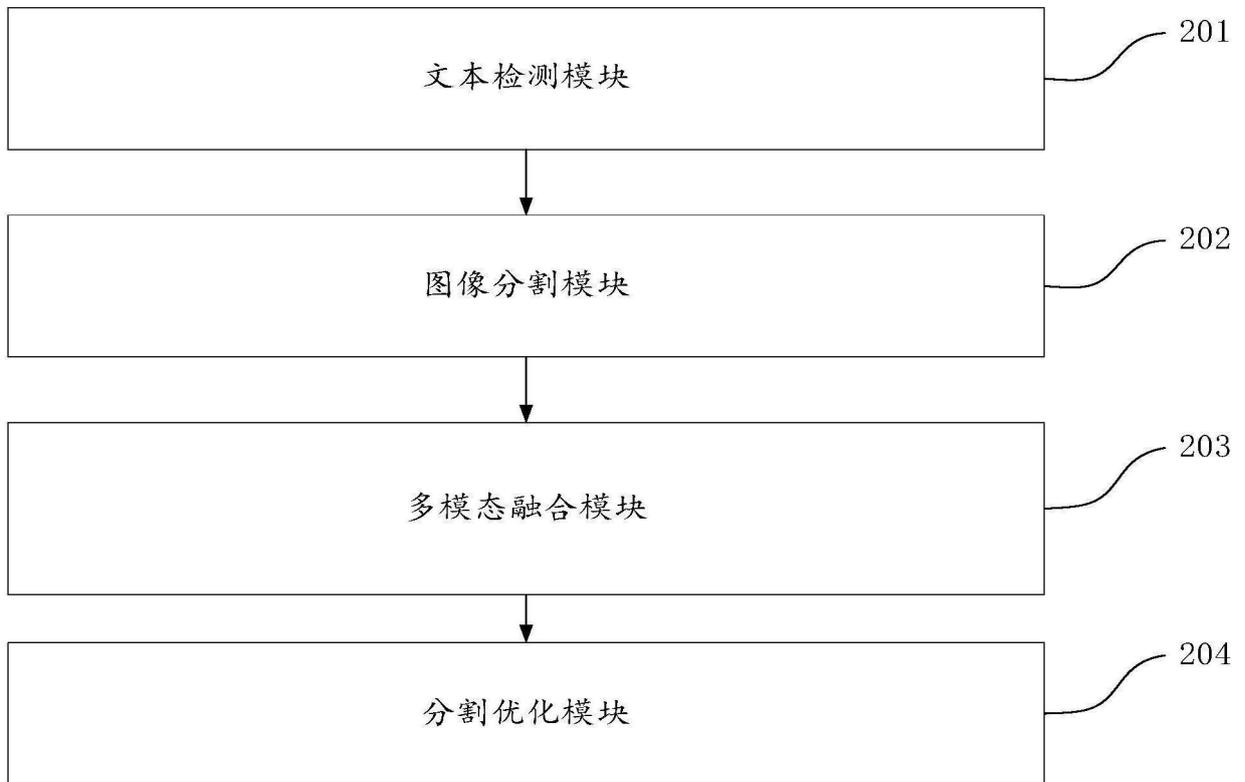


图2

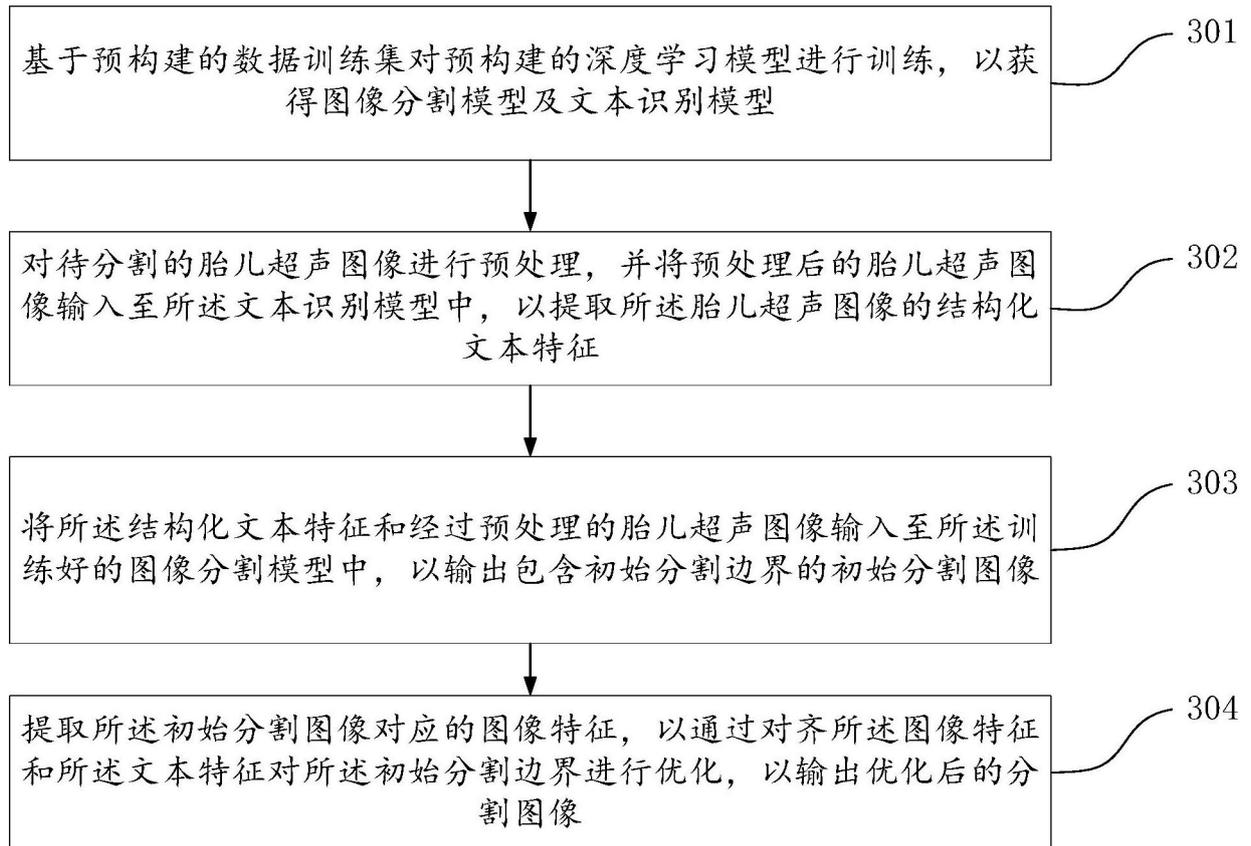


图3

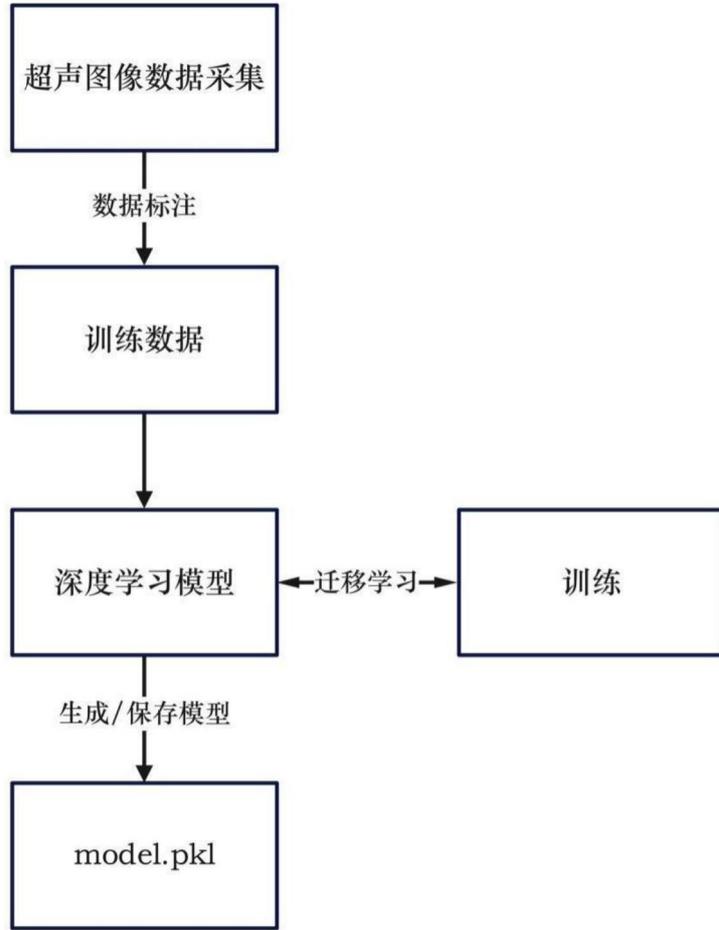


图4

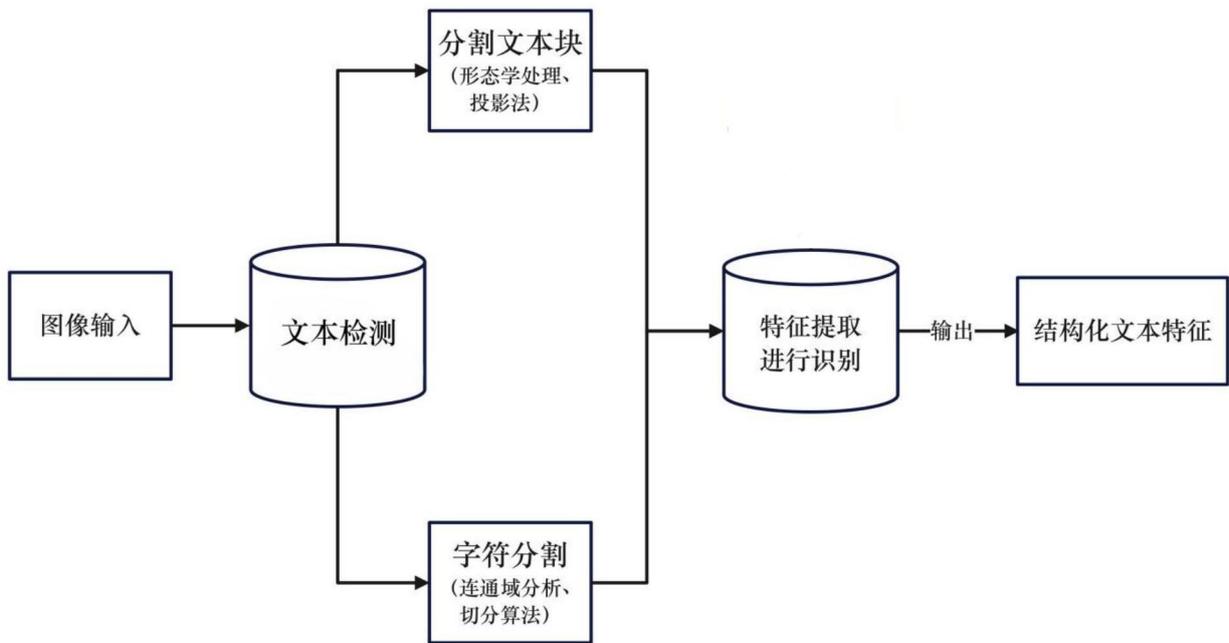


图5

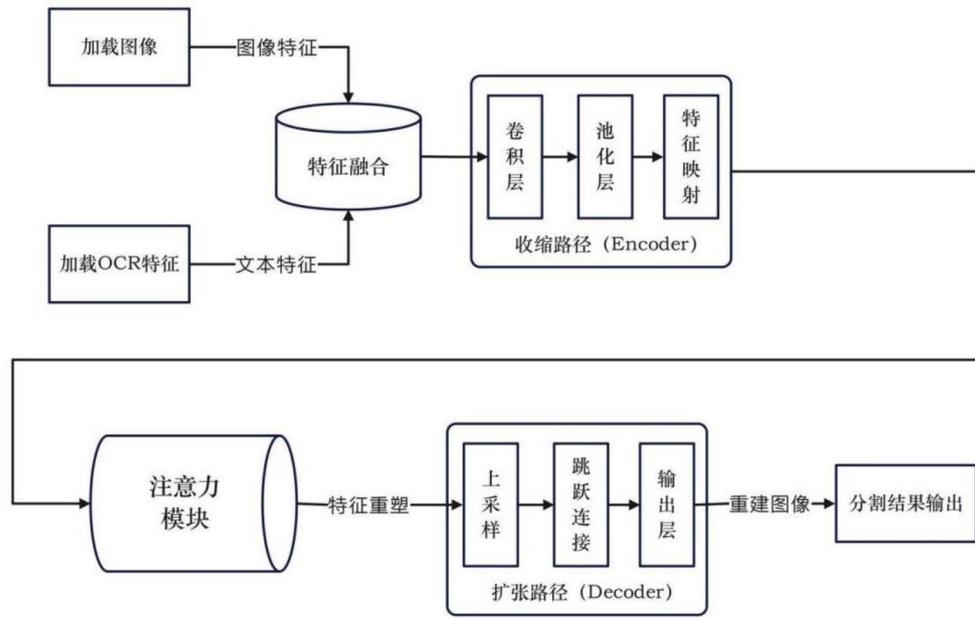


图6

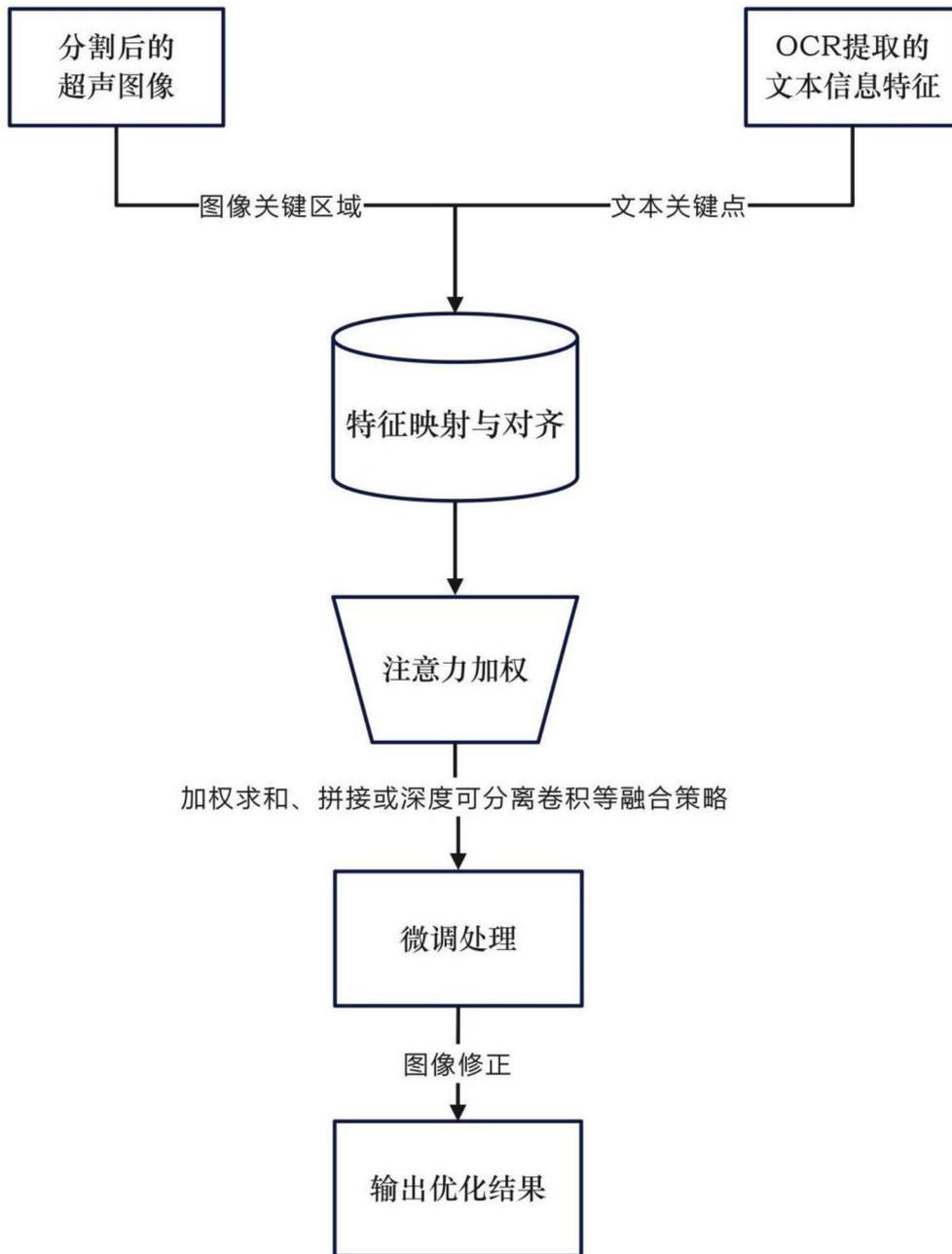


图7