

一种巡查场景的定位方法及系统

本申请要求于 2024 年 10 月 21 日提交中国专利局、申请号为 202411464133.4、发明名称为“一种巡查场景的定位方法及系统”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5 技术领域

本申请涉及智能交通技术领域，特别是涉及一种巡查场景的定位方法及系统。

背景技术

随着基建的高峰过去，城市高架、城市快速路、高速公路、隧道等道路基础设施面临着巨大的管养压力。在服役一定年限后，路面、桥梁结构等设施会在气温、雨水、荷载等多重因素的作用下，会出现各类缺陷病害，轻微10 的病害会影响设施的美观，而严重的病害不仅影响美观，甚至还会影响慢行交通的安全。

要延长道路基础设施的使用性能，最重要的就是要提高发现基础设施上的病害、事件的发现能力，并对病害的定位、尺寸、类型等信息进行准确的15 描述。其中，定位数据，是病害的一个重要信息，养护班组获取定位数据后，可以根据定位找到对应的病害、事件进行维修。此外，定位数据还可帮助进一步的分析应用：利用定位数据可以对病害、事件在空间上的分布进行分析，例如分析病害、事件在不同空间上的聚集密度及分布规律；也可以对同一位置上的病害、事件进行时间、类别上的解析，对同一位置多次上报的病害进行20 空间去重，或分析该处位置是否出现“频修频补”的状态。

目前道路设施的日常巡查，主要是由巡查员在徒步或驾车过程中，通过目测进行巡查，当巡查中发现病害、事件时，巡查员通过手机 APP 对病害、事件进行图像采集，并利用手机的定位功能对该点的定位数据进行记录、25 上报。而随着科技进步，也有部分巡查开始使用专业的巡查设备采集图像，使用人工智能算法识别病害、事件，并使用高精度定位传感器获取准确的位

置信息，减小定位数据的误差，提高定位精度。

目前的日常巡查采集的定位数据通常是二维的经纬度数据，并且可以在电子地图上进行撒点显示，根据经纬度来实现路段位置的匹配。但是在高架-地面道路、隧道-地面、主线-匝道等特殊情况，在同一经纬度上存在上下

5 重叠的情况，单纯使用经纬度，无法判别当前的位置是属于上层还是下层，无法实现准确的定位和路段匹配，进而可能导致识别派单出现问题。目前，既有的一些解决思路、方法包括：（1）、人工进行记录：巡查、检测中，由巡查员观测，并在 APP 上记录当前的巡查场景位置，或在巡查设备上增设开关，由巡查员点击开关，记录不同巡查时间对应的巡查场景。这种方法
10 主要依赖人工的主观判断及操作，容易受巡查员的主管影响导致数据不准确。

（2）、采集轨迹、海拔及加速度数据：在巡查设备上增加高精度导航定位接收机及三轴加速度传感器，对巡查车辆轨迹、速度、海拔、三轴加速度数据进行记录，使用车辆的轨迹及三轴加速度数据，对车辆的行驶路线、上下坡时间点进行标记，根据车辆的海拔高度、行驶状态判断车辆当前可能是在
15 重叠道路的哪一层。这种方法比较依赖传感器数据及基础数据的获取：一方面，当卫星信号不好时，采集的轨迹、速度、海拔等数据有较大的偏差，甚至可能是无效数据，进而导致定位不准确；另一方面，相同的设施，海拔区分阈值是不同的，而高架、隧道的海拔区分阈值较难获取，这些都对准确的场景区分及后续的分析带来干扰。（3）、利用分类神经网络处理数据：部
20 分巡查设备上装配了相机，将相机采集的图像输入给分类神经网络，由神经网络判别图像所属的场景。这种方式虽然能对图像场景做一个判断，但是图像仅做有限类别的分类，无法对不同个体做进一步区分，例如：如果同一桥梁的 2 条桥梁的辅道挨得很近、方向相同，且都属于桥梁辅道，单纯依靠图像分类模型的结果无法区分这两条路，进而导致道路匹配的不准确。

25 **发明内容**

本申请的目的是提供一种巡查场景的定位方法及系统，使用边云结合的系统架构，利用人工智能图像分析算法，实现巡查设备的场景准确、快速定

位，同时利用向量数据库，对道路场景进行特征建档，提升图像场景分类的准确性，有助于后续的进一步数据匹配和分析。

为实现上述目的，本申请提供了如下方案：

5 第一方面，本申请提供一种巡查场景的定位方法，所述巡查场景的定位方法包括：

实时采集巡查场景的图像数据和定位数据；

对所述图像数据进行语义分割和特征提取，获取场景分类结果和特征向量；

10 基于所述图像数据、所述场景分类结果、所述特征向量进行病害识别，若识别出病害，则基于所述定位数据、所述特征向量在预设的场景向量数据库中混合搜索，获取所述巡查场景的混合搜索结果；

将所述混合搜索结果更新到所述场景向量数据库，完成所述巡查场景的病害定位。

15 在一示例性实施例中，对所述图像数据进行语义分割和特征提取，获取场景分类结果和特征向量包括：

将原始图像输入语义分割神经网络模型 M1 中，输出语义分割图像；

将所述语义分割图像与所述原始图像进行组合拼接，获取组合图像；

将所述组合图像输入神经网络模型 M2 中，获取所述场景分类结果和所述特征向量。

20 在一示例性实施例中，所述语义分割神经网络模型 M1 采用图像编码神经网络，包括编码器和解码器；所述编码器用于将所述原始图像提取为高维特征编码向量；所述解码器用于将所述高维特征编码向量解码为与所述原始图像尺寸相同的语义分割图像。

25 在一示例性实施例中，所述神经网络模型 M2 采用卷积神经网络，包括主干网络部分和头部网络部分；所述主干网络部分采用多层卷积和池化层对所述组合图像进行处理，获取高维特征图，再采用全局平均池化层对所述高维特征图进行处理，获取所述特征向量；所述头部网络部分采用包含残差块

的多层神经网络对所述特征向量进行处理，获取所述场景分类结果。

在一示例性实施例中，所述进行混合搜索包括：

根据所述定位数据中的经纬度和方位角，在所述场景向量数据库中进行初步筛选，获取记录数据集合；

- 5 计算所述特征向量和所述记录数据集合的余弦相似度，并基于所述余弦相似度在所述记录数据集合中进行再次筛选，获取混合搜索结果集合。

在一示例性实施例中，将所述混合搜索结果更新到所述场景向量数据库包括：

- 10 若所述混合搜索结果集合为空集，则通过人工对所述图像数据、所述场景分类结果、所述特征向量、所述定位数据进行审核后，添加至所述场景向量数据库；

若所述混合搜索结果集合不为空集，且所述混合搜索结果集合中的个体ID和场景类别都相同，则将所述图像数据、所述场景分类结果、所述特征向量、所述定位数据更新至所述场景向量数据库；

- 15 若所述混合搜索结果集合不为空集，且所述混合搜索结果集合中的个体ID存在不同，则将所述混合搜索结果集合中的特征向量进行算数平均，获取新特征向量并重新在所述场景向量数据库中进行混合搜索，获取新混合搜索结果集合并进行投票，获取修正后的个体ID和场景类别并更新至所述场景向量数据库。

- 20 在一示例性实施例中，将所述混合搜索结果更新到所述场景向量数据库后，还包括通过人工对所述场景向量数据库进行定期清洗审核。

- 第二方面，本申请还提供了一种巡查场景的定位系统，所述巡查场景的定位系统包括：边缘端设备和云端设备，其中，所述边缘端设备用于实时采集巡查场景的图像数据和定位数据，并对所述图像数据进行特征提取，获取
25 场景分类结果和特征向量；所述云端设备用于对所述图像数据、所述定位数据、所述场景分类结果和特征向量进行处理，获取所述巡查场景的病害结果和病害定位。

在一示例性实施例中，所述边缘端设备包括工控机、高清相机和导航定位接收机；所述高清相机和所述导航定位接收机连接到所述工控机上；所述工控机用于提供数据传输、图像特征提取以及供电功能。

在一示例性实施例中，所述云端设备包括病害识别模块、混合搜索模块、
5 派单模块；所述病害识别模块用于基于所述图像数据、所述场景分类结果、
所述特征向量进行病害识别；所述混合搜索模块用于当识别出病害时，基于
所述定位数据、所述特征向量在预设的场景向量数据库中进行混合搜索，获
取所述巡查场景的混合搜索结果；所述派单模块用于将所述混合搜索结果更
新到所述场景向量数据库并结合病害识别结果发送给相应养护人员。

10 根据本申请提供的具体实施例，本申请公开了以下技术效果：

（1）边缘计算与云计算相结合，提升效率：采用边云结合的方式，无
需人工现场进行记录、测量，使得实时数据处理更加高效、快速。边缘端设
备负责现场的数据收集和初步处理，减少大量原始数据传输的压力，同时确
保了数据的安全性和隐私保护。

15 （2）高精度图像识别及人工校验互补：通过集成高清相机和智能算法，
对图像中的交通设施进行语义目标提取，并将提取结果与原始图像结合，让
后续的分析模型获得更多现场基础设施的信息。同时本申请保留了人工审核
环节，确保了结果的准确性，特别是在面对不确定或模糊的情况时，人工干
预可以提供最终决定。

20 （3）多层次数据分析：在进行场景及 ID 匹配时，使用了混合搜索，充
分利用了采集的经纬度、方位角、特征向量数据等多元数据，有助于提升搜
索速度及精度；相对于单纯使用分类判别，本申请考虑了同类别不同个体存
在的差异性，对于物理距离较近，场景类别相同，但属于不同设施的情况，
可以进行有效的区分。

25 （4）数据库优化管理：本申请采用的向量数据库，是一种复合式关系
型数据库，支持高效的查询和存储功能，使用时不需要记录特殊的字段阈值
进行分类，便于管理和维护大量的历史数据，同时也提供了灵活的数据更新

机制，适应不断变化的道路情况。

附图说明

为了更清楚地说明本申请实施例或相关技术中的技术方案，下面将对实施例中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 为本申请实施例的一种巡查场景的定位方法流程图；

图 2 为本申请实施例的智能巡查设备安装示意图；

图 3 为本申请实施例的采集图像的灭点高度示意图；

10 图 4 为本申请实施例的数据处理流程图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请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申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，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，都属于本申请保护的范围。

本申请的目的是提供一种巡查场景的定位方法及系统，使用边云结合的系统架构，利用人工智能图像分析算法，实现巡查设备的场景准确、快速定位，同时利用向量数据库，对道路场景进行特征建档，提升图像场景分类的准确性，有助于后续的进一步数据匹配和分析。

20 为使本申请的上述目的、特征和优点能够更加明显易懂，下面结合附图和具体实施方式对本申请作进一步详细的说明。

本实施例提供了一种巡查场景的定位方法，如图 1 所示，包括：

实时采集巡查场景的图像数据和定位数据；

对图像数据进行语义分割和特征提取，获取场景分类结果和特征向量；

25 基于图像数据、场景分类结果、特征向量进行病害识别，若识别出病害，则基于定位数据、特征向量在预设的场景向量数据库中进行混合搜索，获取巡查场景的混合搜索结果；

将混合搜索结果更新到场景向量数据库，完成巡查场景的病害定位。

具体的，本实施例实时采集巡查场景的图像数据和定位数据，通过在巡查汽车上安装智能巡查装备实现，且智能巡查装备上配置智能算法，能对图像进行语义分割、特征提取、图像分类，并将采集分析的数据上传至云端。

- 5 云端在接收到采集的数据后，执行特征向量混合搜索，并根据后处理得到对应的分析结果，将相关数据入库保存。

智能巡查装备组成及安装如图 2 所示，包含 1 个及以上的高清相机以及一个工控机，其中工控机上还连接一个高精度导航定位接收机，有互联网卡，并有一定的 AI 算力，可以部署识别算法程序；供电来源上，智能巡查装备
10 使用直流电供电；防水性能方面，智能巡查装备具有良好的防水性能。

工控机安装于车辆内部，为车载供电。其具有第四代(4G)或第五代(5G)移动通信技术，连接高精度导航定位接收机和高清相机。可以向高清相机发送采集图像的指令，并不间断地接受、处理、上传采集的数据。此外，中央工控机内也可以包含微型图像处理器（Graphics processing unit, GPU），深度神经网络算法可利用 GPU 加速图像处理。
15

高清相机，通过工控机或车载供电，相机像素 ≥ 100 万像素，采集时画质清晰，无延迟，采集帧率 $\geq 25\text{Hz}$ ，其可安装于车内或者车外，安装后视角固定，可用于采集车辆前方或车辆后方图像。采集的图像要求画质清晰，无明显畸变，没有被明显遮挡，图像中灭点到图像上边缘的高度 h' 与图像总高度 h 的比值，如图 3 所示，要求限定在一定范围，即 $1/3 \leq h'/h \leq 3/4$ 。
20

图像灭点可通过在工控机上部署一个卷积神经网络分类模型，将图像竖向等分为 n 个区间，并预测图像灭点在哪个区间的概率最大。

高精度导航定位接收机的定位误差 $\leq 1\text{m}$ ，并配置有惯性导航系统，定位数据的采集频次 $\geq 10\text{Hz}$ 。

- 25 进一步地，本实施中数据的识别及处理，使用的是边缘端设备感知识别+云端数据后处理执行，处理流程及逻辑如图 4 所示。

对图像数据进行语义分割和特征提取，获取场景分类结果和特征向量包

括：将原始图像输入语义分割神经网络模型 M1 中，输出语义分割图像；将语义分割图像与原始图像进行组合拼接，获取组合图像；将组合图像输入神经网络模型 M2 中，获取场景分类结果和特征向量。

语义分割神经网络模型 M1 采用图像编码神经网络，包括编码器和解码器，编码器用于将原始图像提取为高维特征编码向量，解码器用于将高维特征编码向量解码为与原始图像尺寸相同的语义分割图像。

神经网络模型 M2 采用卷积神经网络，包括主干网络部分和头部网络部分，主干网络部分采用多层卷积和池化层对组合图像进行处理，获取高维特征图，再采用全局平均池化层对高维特征图进行处理，获取特征向量；头部网络部分采用包含残差块的多层神经网络对特征向量进行处理，获取场景分类结果。

进行混合搜索包括：根据定位数据中的经纬度和方位角，在场景向量数据库中初步筛选，获取记录数据集合；计算特征向量和记录数据集合的余弦相似度，并基于余弦相似度在记录数据集合中进行再次筛选，获取混合搜索结果集合。

将混合搜索结果更新到场景向量数据库包括：

若混合搜索结果集合为空集，则通过人工对图像数据、场景分类结果、特征向量、定位数据进行审核后，添加至场景向量数据库；

若混合搜索结果集合不为空集，且混合搜索结果集合中的个体 ID 和场景类别都相同，则将图像数据、场景分类结果、特征向量、定位数据更新至场景向量数据库；

若混合搜索结果集合不为空集，且混合搜索结果集合中的个体 ID 存在不同，则将混合搜索结果集合中的特征向量进行算数平均，获取新特征向量并重新在场景向量数据库中进行混合搜索，获取新混合搜索结果集合并进行投票，获取修正后的个体 ID 和场景类别并更新至场景向量数据库。

具体的，本实施例的巡查汽车在执行巡查时，采集原始图像 m_i ，并使用语义分割神经网络模型 M1，对图像进行语义分割，获得语义分割图像 m_i' ；

再将原始图像 m_i 与语义分割图像 m_i' 进行组合拼接, 得到整合后的图像 m_i^* ; 将整合后的组合图像 m_i^* 输入到神经网络模型 M2, 获得场景分类结果 c_i , 并输出一个对应的特征向量 v_i ; 在完成图像处理后, 将采集图像、 c_i 、 v_i 及对应的经纬度、方位角数据上传到云端服务器。

- 5 原始图像 m_i , 图像尺寸一般为 $[u, v, 3]$, 即包含 RGB 三个通道, 图像的长宽比例一般为 16:9 或 4:3。

方位角, 也可成为航向角, 用于采集图像对于高清相机的一个参考方向, 以正北方向为 0° , 顺时针方向增大, 数据取值范围是 $[0, 360)$ 。

- 语义分割神经网络模型 M1, 是指一种图像编解码类型的神经网络, 先
10 通过特征提取编码器神经网络层, 将原始图像 m_i 提取成一个高维度的特征编码向量, 再通过特征解码器神经网络层, 将特征向量解码成一个与原图像素尺寸相同的语义分割图像 m_i' 。语义分割主要将图像中的路面区域、护栏等交安设施、标牌等路侧设施、跨线桥梁、高架结构等道路设施进行类别标记, 并根据类别分配一个 $[100, 255]$ 中一个指定的数值, 其他未被标记的区域
15 统一归类为 background 类别, 并且标记为 0, 实现对图像里所有像素点的类别标记, 进而生成语义分割图像 m_i' , 图像尺寸为 $[u, v]$ 。根据类别分配一个 $[100, 255]$ 中一个指定的数值, 是指从 $[100, 255]$ 中挑选 j 个整数, 对应于图像语义识别网络中能识别的 j 种设施, 按照一个类别一个整数来进行对应。

- 原始图像 m_i 与语义图像 m_i' 进行组合拼接, 是指将两个图像按照通道方向
20 进行组合, 形成新的组合图像 m_i^* , 图像尺寸为 $[u, v, 4]$, 4 个通道分别为 RGB 图像及语义分割图像。

神经网络模型 M2 由两部分组成: 主干网络部分, 是卷积神经网络, 按照 Re-Identification (ReID) 进行训练; 在主干网络的基础上, 增加一个全连接分类头网络, 用于类别分类。

- 25 1) 主干网络部分: 模型先将组合图像 m_i^* 进行缩小、归一化等预处理, 使用多层卷积 (CNN) + 池化 (Pooling) 网络对组合图像 m_i^* 进行处理, 得到高维特征图 $[u', v', k]$, 再使用全局平均池化对高维特征图进行处理, 得

到 512 维度的特征向量 v_i 。

2) 头部网络部分: 将 v_i 输入到一个包含残差块的多层神经网络中进行进一步处理, 并在神经网络最后输出 k 个分类及对应的置信度, 其中 k 个分类对应于模型要进行的 k 个场景分类, 包括但不限于“地面道路”、“快速
5 路”、“辅道”、“隧道”、“高架下方道路”等类别, 并以置信度最大的类别作为 M2 模型判别的场景分类结果, 记为 c_i 。

在完成分析后, 工控机将 m_i 、 c_i 、 v_i , 以及时间对应的经纬度 $[lng_i, lat_i]$ 、方位角 azi_i , 一并上传到云端服务器。其中, 经纬度 $[lng_i, lat_i]$ 的数值精度达 0.000001。

10 云端服务器在接收到新的采集数据后, 会使用经纬度数据 $[lng_i, lat_i]$ 、方位角数据 azi_i 、特征向量 v_i 、分类类别 c_i 数据, 在云端的场景向量数据库中, 进行混合搜索, 根据混合搜索的结果, 再执行对应的处理逻辑, 完成采集图像的场景判断, 并对云端场景向量数据库进行更新。

场景向量数据库, 是一种复合式关系型数据库, 包含多条记录, 数据库
15 内的每一条记录, 包含了一个采集图像的链接 url 、该图像的经纬度 $[lng, lat]$, 采集该图像的方位角 azi , 该图像的场景分类结果 c , 该图像的个体 ID, 该图像经过特征提取之后的特征向量 v , 该记录的创建时间 $CREATETIME$, 是否经过审核员校验 $CHECK$ 及是否标记删除 $DELETE$ 。向量数据库内的数据是动态变化的, 完成校核后的数据可以添加至向量数据库内, 审核员也可以
20 定期对向量数据库进行清洗校核, 校正图像的场景类别。

混合搜索, 是指根据上传数据的经纬度及方位角, 对数据库中的数据
进行初步筛选, 筛选出 $[lng_i, lat_i]$ 附近, 方位角与 azi_i 相近的记录数据集 $Vset_i$, 再使用余弦距离, 计算特征向量 v_i 和 $Vset_i$ 的相似度, 设置余弦相似度的阈值 T , 从 $Vset_i$ 中挑选相似度 $\geq T$ 的新的向量集合, 作为混合搜索结果集合
25 $Vset_i'$ 。

根据混合搜索结果集合 $Vset_i'$ 的情况, 进行对应的业务后处理:

1) 、 $Vset_i'$ 为空集, 即表示向量数据库在经纬度 $[lng_i, lat_i]$ 附近, 没有与

v_i 相似的向量,则将设备上传的 v_i 、 c_i 、 m_i 等数据推送至待确认队列中,经审核员校核确认后,添加至向量数据库 Vdb 中。

2)、如果 $Vset_i'$ 不为空集,且 $Vset_i'$ 的个体 ID 都是相同的,场景类别都为 c_i ,则判定采集的数据对应的场景分类结果为 c_i ,并将 v_i 、 c_i 及个体 ID 等数据更新追加到 Vdb 中。

3)、如果 $Vset_i'$ 不为空集,且 $Vset_i'$ 的个体 ID 存在不相同,则将 $Vset_i'$ 中的特征向量进行算数平均,获得新的特征向量 v'' ,再使用 $[lng_i, lat_i]$ 、 azi_i 、 v'' 去 Vdb 中进行新的特征检索,获取相似度最高的 e 个向量,组成集合得到新的特征向量集合 $Vset_i''$,根据向量对应的个体 ID 进行投票,以投票的结果中最可能的个体 ID 对应的场景类别 c'' 作为采集图像的场景类别,并将 c'' 、 $[lng_i, lat_i]$ 、 azi_i 、最可能的个体 ID 等数据更新到 Vdb 中。

投票方式包括但不限于:

- i、多数投票:以 e 个向量中,以数量统计最多的场景类别作为投票结果。
 - ii、加权投票:根据余弦距离加权投票,距离越小,可投的票数越多,
- 最后统计 e 个向量中,哪个个体 ID 的投票数量最多。

进一步地,将混合搜索结果更新到场景向量数据库后,还包括通过人工对场景向量数据库进行定期清洗审核。

具体的,审核员可以使用辅助工具,定期对既有的向量数据库 Vdb 进行清洗,包括但不限于以下的方法:

1) 使用区别于神经网络模型 M2 的新的神经网络模型 $M2'$,对 Vdb 的数据进行重新识别,并挑选出识别结果与 Vdb 中不同的数据记录,推送至待确认队列,由审核员进行校核,对不匹配的场景类别进行修正,并更新到数据库内。

2) 将 Vdb 中个体 ID 相同的图像提取出来,并将不属于该 ID 的记录数据进行清洗。

为进一步优化技术方案,本实施例还提供了一种巡查场景的定位系统,包括:边缘端设备和云端设备,边缘端设备用于实时采集巡查场景的图像数

据和定位数据，并对图像数据进行特征提取，获取场景分类结果和特征向量；云端设备用于对图像数据、定位数据、场景分类结果和特征向量进行处理，获取巡查场景的病害结果和病害定位。

进一步地，边缘端设备包括工控机、高清相机和导航定位接收机，高清相机和导航定位接收机连接到工控机上，工控机提供数据传输、图像特征提取以及供电功能。

具体的，边缘端设备主要为车载式智能巡查设备，设备相关的描述如下：

1) 配备工控机，工控机上配备 NVIDIA GT1030 显卡，能够运行深度学习算法程序。工控机采用直流电源供电，确保稳定运行。工控机具有良好的防水性能，适应各种户外条件。

2) 工控机内置高精度 GPS 模块和互联网卡，可采集经纬度及方位角，采集帧率为 10Hz，支持实时向云端服务器进行数据传输。

3) 在车顶上方安装有一个高清相机连接到工控机，相机采集图像分辨率为 200 万像素，图像采集帧率为 25Hz，图像清晰、无遮挡，相机朝车辆正前方，相机视角水平。

边缘端设备数据采集及数据处理的流程如下：

1) 边缘端设备在通电后，会从 GPS 传感器采集 gps 数据，根据 gps 状态判断当前车辆是否在巡查，当车辆处于移动状态时，工控机会对车辆的位移进行一个距离的累计，当累计距离达到 10m 时，工控机下发一个采集图像和 gps 数据的指令，并从 0 开始进行新的距离累计。

2) 采集后的图像 m_i ，会输入到一个语义分割神经网络模型 M1，该模型主要是对图像中的交安设施、路面区域、桥梁高架等设施进行语义感知，进行图像处理，生成一个语义分割图像 m_i' ，像素尺寸与原图 m_i 相同， m_i' 是对图像范围内的每个像素点分配一个类别，包括护栏、防眩板、路面、桥梁、交通标牌、天空、警示杆、background 等，并根据类别分配一个具体[0, 255) 的整数。

3) 将图像 m_i 与语义分割图像 m_i' ，按照通道进行拼接，获得新的组合

图像 m_i^* , 将 m_i^* 进行缩放, 在输入到神经网络模型 M2, 该模型会输出一个 512 维度的向量 v_i , 并会输出一个 k 维度的置信度数组, 对应于“快速路”、“地面道路”、“隧道”等 k 个场景类别, 其中置信度最高的则为模型推理的场景类别, 记录为 c_i 。

- 5 4) 将上述采集的图像数据 m_i 、经纬度数据 $[lng_i, lat_i]$ 、方位角 azi_i 以及分析的结果数据 v_i 、 c_i 进行汇总, 上传至云端设备。

进一步地, 云端设备包括病害识别模块、混合搜索模块、后处理模块, 病害识别模块用于基于图像数据、场景分类结果、特征向量进行病害识别; 混合搜索模块用于当识别出病害时, 基于定位数据、特征向量在预设的场景
10 向量数据库中进行混合搜索, 获取巡查场景的混合搜索结果; 后处理模块用于将混合搜索结果更新到场景向量数据库并结合病害识别结果发送给相应养护人员。

具体的, 云端设备主要为云端服务器, 服务器上安装有数据接收程序、场景向量数据库、数据处理分析程序、派单程序等。

- 15 1) 病害识别模块:

云端服务器在接收到新的数据, 包含了图像数据 m_i 、经纬度数据 $[lng_i, lat_i]$ 、方位角 azi_i 以及分析的结果数据 v_i 、 c_i , 会将图像数据输入到病害识别模型中, 如果识别出图像有病害, 则将采集数据发送给后处理程序进行进一步分析, 本实施例中病害识别模型采用目标识别神经网络, 能够在图像上标
20 记出识别的病害类别和识别框位置等信息。

- 2) 混合搜索模块:

接到数据后, 会使用经纬度 $[lng_i, lat_i]$ 、方位角数据 azi_i 、特征向量数据 v_i , 在向量数据库 MILVUS 中进行混合搜索。MILVUS 向量数据库包含的字段信息有:

- 25 lng : 经度;
 lat : 纬度;
 azi : 方位角, 取值范围是 $[0, 360)$;

vec: 特征向量;

CHECK: 审核员是否进行了校验, 1 为经过校核, 0 为未经过校核;

DELETE: 数据是否标记删除, 1 为标记删除, 0 为未标记删除;

CREATETIME: 采集时间;

5 c: 场景类别;

ID: 个体 ID。

其中, 进行混合搜索时, 主要遵循以下搜索条件:

i、 $\text{lng}_i - 0.00001 * 8 < \text{lng}_i + 0.00001 * 8$;

ii、 $\text{lat}_i - 0.00001 * 8 < \text{lat}_i + 0.00001 * 8$;

10 iii、 $\text{azi}_i - 5 < \text{azi} < \text{azi}_i + 5$ ($\text{azi}_i \in [5, 355]$);

$0 < \text{azi} < \text{azi}_i + 5 \cup \text{azi} > \text{azi}_i - 5 + 360$ ($\text{azi}_i \in [0, 5)$);

$\text{azi}_i - 5 < \text{azi} < 360 \cup \text{azi} < \text{azi}_i + 5 - 360$ ($\text{azi}_i \in (355, 360)$);

iv、DELETE=0, CHECK=1;

v、搜索得到的向量, 余弦距离>0.8。

15 根据混合搜索的限制条件, 得到搜索结果集合 $V\text{set}_i$, 对搜索结果集合 $V\text{set}_i$ 进行分析包括:

i、 $V\text{set}_i'$ 为空集, 即表示向量数据库在经纬度 $[\text{lng}_i, \text{lat}_i]$ 附近, 没有与 v_i 相似的向量, 则将设备上传的 v_i 、 c_i 、 m_i 等数据推送至待确认队列中, 经审核员校核确认后, 添加至向量数据库 $V\text{db}$ 中。

20 ii、如果 $V\text{set}_i'$ 不为空集, 且 $V\text{set}_i'$ 的个体 ID 都是相同的, 场景类别都为 c_i , 则判定采集的数据对应的场景分类结果为 c_i , 并将 v_i 、 c_i 及个体 ID 等数据更新追加到 $V\text{db}$ 中。

iii、如果 $V\text{set}_i'$ 不为空集, 且 $V\text{set}_i'$ 的个体 ID 存在不相同, 则将 $V\text{set}_i'$ 中的特征向量进行算数平均, 获得新的特征向量 v'' , 再使用 $[\text{lng}_i, \text{lat}_i]$ 、 azi_i 、 v'' 去 $V\text{db}$ 中进行新的特征检索, 获取相似度最高的 10 个向量, 组成集合得到新的特征向量集合 $V\text{set}_i''$, 根据向量对应的个体 ID 进行投票, 以投票的结果中最可能的个体 ID 对应的场景类别 c'' 作为采集图像的场景类别, 并将

25

c'' 、 $[lng_i, lat_i]$ 、 azi_i 、最可能的个体 ID 等数据更新到 Vdb 中。使用的投票逻辑是，根据余弦距离加权投票，距离越小，可投的票数越多，可投票数依次为 $[10, 9, 8, 7, 6, 5, 4, 3, 2, 1]$ ，最后统计 10 个向量中，哪个个体 ID 的投票数量最多。

5 3) 派单模块:

经过后处理的数据，如果不需要进行审核员校核，将直接推送给派单模块；如果需要进行审核员校核，待审核员校核通过后，相关的数据会推送到派单模块。在派单模块内，病害数据将通过场景类别和设施个体 ID，实现路段的匹配，再推送给具体养护班组，并进行入库归档，以便后续的进一步

10 分析。

进一步地，审核员可以使用辅助工具，定期对既有的场景向量数据库 Vdb 进行清洗，包括以下的方法:

1) 深度更深的神经网络模型 $M2'$ ，对 Vdb 的图像数据进行重新识别，并挑选出识别结果与 Vdb 中不同的数据记录，推送至待确认队列，由审核员

15 进行校核，对不匹配的场景类别进行修正，并更新到数据库内。

2) 将 Vdb 中个体 ID 相同的图像提取出来，并将不属于该 ID 的记录数据进行清洗。

本说明书中各个实施例采用递进的方式描述，每个实施例重点说明的都是与其他实施例的不同之处，各个实施例之间相同相似部分互相参见即可。

20 以上所述的实施例仅是对本申请优选方式进行的描述，并非对本申请的范围进行限定，在不脱离本申请设计精神的前提下，本领域普通技术人员对本申请的技术方案做出的各种变形和改进，均应落入本申请权利要求书确定的保护范围内。

权 利 要 求 书

1.一种巡查场景的定位方法，其特征在于，所述巡查场景的定位方法包括：

实时采集巡查场景的图像数据和定位数据；

对所述图像数据进行语义分割和特征提取，获取场景分类结果和特征向量；

基于所述图像数据、所述场景分类结果、所述特征向量进行病害识别，若识别出病害，则基于所述定位数据、所述特征向量在预设的场景向量数据库中

5 量；
将所述混合搜索结果更新到所述场景向量数据库，完成所述巡查场景的病害定位。

2.根据权利要求1所述的巡查场景的定位方法，其特征在于，对所述图像数据进行语义分割和特征提取，获取场景分类结果和特征向量包括：

将原始图像输入语义分割神经网络模型M1中，输出语义分割图像；

将所述语义分割图像与所述原始图像进行组合拼接，获取组合图像；

15 将所述组合图像输入神经网络模型M2中，获取所述场景分类结果和所述特征向量。

3.根据权利要求2所述的巡查场景的定位方法，其特征在于，所述语义分割神经网络模型M1采用图像编码神经网络，包括编码器和解码器；所述编码器用于将所述原始图像提取为高维特征编码向量；所述解码器用于将所述高维特征编码向量解码为与所述原始图像尺寸相同的语义分割图像。

4.根据权利要求2所述的巡查场景的定位方法，其特征在于，所述神经网络模型M2采用卷积神经网络，包括主干网络部分和头部网络部分；所述主干网络部分采用多层卷积和池化层对所述组合图像进行处理，获取高维特征图，再采用全局平均池化层对所述高维特征图进行处理，获取所述特征向量；所述头部网络部分采用包含残差块的多层神经网络对所述特征向量进行处理，获取所述场景分类结果。

5.根据权利要求1所述的巡查场景的定位方法，其特征在于，所述进行

混合搜索包括:

根据所述定位数据中的经纬度和方位角,在所述场景向量数据库中进行初步筛选,获取记录数据集合;

5 计算所述特征向量和所述记录数据集合的余弦相似度,并基于所述余弦相似度在所述记录数据集合中进行再次筛选,获取混合搜索结果集合。

6.根据权利要求 5 所述的巡查场景的定位方法,其特征在于,将所述混合搜索结果更新到所述场景向量数据库包括:

10 若所述混合搜索结果集合为空集,则通过人工对所述图像数据、所述场景分类结果、所述特征向量、所述定位数据进行审核后,添加至所述场景向量数据库;

若所述混合搜索结果集合不为空集,且所述混合搜索结果集合中的个体 ID 和场景类别都相同,则将所述图像数据、所述场景分类结果、所述特征向量、所述定位数据更新至所述场景向量数据库;

15 若所述混合搜索结果集合不为空集,且所述混合搜索结果集合中的个体 ID 存在不同,则将所述混合搜索结果集合中的特征向量进行算数平均,获取新特征向量并重新在所述场景向量数据库中进行混合搜索,获取新混合搜索结果集合并进行投票,获取修正后的个体 ID 和场景类别并更新至所述场景向量数据库。

20 7.根据权利要求 1-6 任一项所述的巡查场景的定位方法,其特征在于,将所述混合搜索结果更新到所述场景向量数据库后,还包括通过人工对所述场景向量数据库进行定期清洗审核。

25 8.一种巡查场景的定位系统,其特征在于,所述巡查场景的定位系统包括:边缘端设备和云端设备,其中,所述边缘端设备用于实时采集巡查场景的图像数据和定位数据,并对所述图像数据进行特征提取,获取场景分类结果和特征向量;所述云端设备用于对所述图像数据、所述定位数据、所述场景分类结果和所述特征向量进行处理,获取所述巡查场景的病害结果和病害定位。

9.根据权利要求 8 所述的巡查场景的定位系统，其特征在于，所述边缘端设备包括工控机、高清相机和导航定位接收机；所述高清相机和所述导航定位接收机连接到所述工控机上；所述工控机用于提供数据传输、图像特征提取以及供电功能。

- 5 10.根据权利要求 8 所述的巡查场景的定位系统，其特征在于，所述云端设备包括病害识别模块、混合搜索模块、派单模块；所述病害识别模块用于基于所述图像数据、所述场景分类结果、所述特征向量进行病害识别；所述混合搜索模块用于当识别出病害时，基于所述定位数据、所述特征向量在预设的场景向量数据库中进行混合搜索，获取所述巡查场景的混合搜索结果；
- 10 所述派单模块用于将所述混合搜索结果更新到所述场景向量数据库并结合病害识别结果发送给相应养护人员。