

洪湖水质净化厂“视频AI智能巡检”

严 勇¹ 刘海牙² 马祥温³

1. 深圳市水务(集团)有限公司 广东 深圳 518000

2. 深圳市水务(集团)有限公司 广东 深圳 518000

3. 深圳市微妙物联科技有限公司 广东 深圳 518000

摘要: 洪湖水质净化厂位于深圳市罗湖区笋岗街道笋岗滞洪区内,为三层全地下式地理污水处理厂,地面为市政公园。深圳毗邻南海,属亚热带海洋性气候,暴雨发生频率高,短时强度大,危害性广,为保证地下生产设备设施正常运行,需要每隔四个小时进行厂区全面巡查。由于为全地下式封闭空间,频繁、长时间的机械式人工巡检,容易激发人固有的不专注性、疲劳性,降低巡检质量;同时如果由于渗水、设备故障、河水倒灌未及时发现,将引发严重安全责任事故。厂区已经应用了国内先进的5G网络、物联网设备、大数据技术。安防系统汇集了海量视频信号,唯独缺少“可以自主做出判断”的人工智能系统进行甄别。针对以上场景,我们设计了一套视频AI智慧巡检系统,实现7*24小时实时监控,视频分析,无人巡检,智能预警。

关键词: 人工智能;深度学习;污水厂巡检

AI巡检包括下面几项:

(1) 人员安全相关: 基于视觉的人员定位分布,安全帽佩戴检测,应急手势识别,人员跌倒识别,抽烟识别,人员闯入危险区域识别。

(2) 生产安全相关: 火焰识别,烟雾识别,水浸识别,水槽异物堵塞识别,泥浆漏泥识别,垃圾桶漫溢识别等。

(3) 生产工艺相关: 设备屏幕读数,状态灯读取,水位监测,设备异响监测,膜池冒泡识别,设备包裹材料脱落识别,链条震动识别等

针对当前的水厂的AI巡检场景,采用摄像头采集监控图像,实现智能巡检的过程中,传统算法识别精度不高,识别效率低的问题,我们提出一种基于深度学习的计算机视觉算法改进算法的图像识别方法。该方法对传统的图像识别算法进行优化,同时改进和优化卷积神经网络,通过多个深度学习模型的级联,增强实际场景中模型的泛化能力,大大提升识别的准确率和不同场景下的泛化性,使得新模型和算法在不同的场景下具备更好通用性。简单描述了算法实现的原理、基于深度学习卷积模型的结构和实现流程,并开发了对应的算法,通过在洪湖水厂智能化AI巡检项目中实际应用,验证结果表明该算法具有良好适应性和高识别可靠度等优点。

1 平台架构

采用人工智能算法为主,和传统算法相结合,运用机器视觉、图像处理、模式识别、目标检测等技术,构建视频智能分析平台,实时监控生产状态和人员安全。

包括: 巡检任务分配,数据标注,训练模型,识别目标等几大模块实现智能化管控,智能运算,视频自动分析,对各类生产巡检内容的监控有效实现智能化监控管理。

项目总体架构如图1:

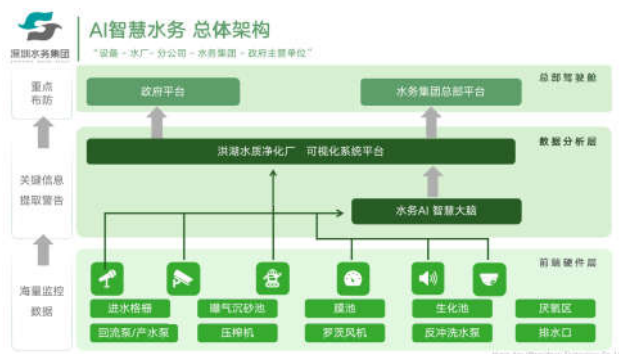


图1: AI智慧巡检总体架构图

2 关键技术及应用

广泛采用了卷积神经网络,目标检测和目标分类模型,并且使用了模型级联的整合模型,达到生产环境下使识别效率达到SOTA的级别。

另外针对声纹和振动波形,我们采用了声纹识别(Voiceprint Recognition)算法,算法模型的基座同样采用了卷积神经网络,声纹特征提取采用了基于傅立叶变换的音频梅尔频谱(mel spectrogram),会把把语音或者震动信号(signa)l变成图片(spectrogram)的形式,然后通过卷积神经网络再做分类或者特征提取。本次项目中采用了ecapa_tdn的处理声纹特征的卷积神经网络模

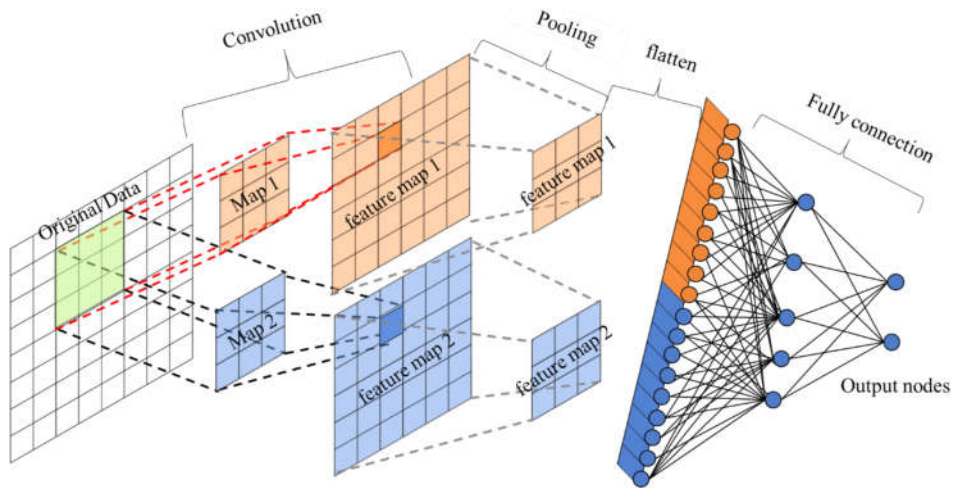
型，ECAPA-TDNN由比利时哥特大学Desplanques等人于2020年提出，通过引入SE (squeeze-excitation)模块以及通道注意机制。

卷积神经网络

2.1 卷积神经网络结构

卷积神经网络 (Convolutional Neural Networks, CNN) 是一类包含卷积计算的前馈神经网络，是基于图

像任务的平移不变性 (图像识别的对象在不同位置有相同的含义) 设计的，擅长应用于图像处理等任务。图像数据具有非常高的维数，卷积神经网络利用了卷积和池化层，能够高效提取图像的重要“特征”，再通过后面的全连接层处理“压缩的图像信息”及输出结果。对比标准的全连接网络，卷积神经网络的模型参数大大减少了。



选图2：典型的卷积神经网络

目标检测算法

2.2 目标检测中边界框回归算法(bounding box regression)

- 训练一个端到端的物体检测器 yolov4
- 构建可以输出边界框坐标的CNN架构，从而使我们实际上可以训练模型以做出更好的对象检测器预测
- 训练这样的模型，通过边界框回归算法达到目标检测

我们都可能熟悉通过深度神经网络进行图像分类的概念。在执行图像分类时，我们：

- 将输入图像呈现给CNN网络
- 通过CNN前馈网络进行预测
- 输出带有N个元素的向量，其中N是类别标签的总数
- 选择可能性最大的类别标签作为我们最终的预测类别标签

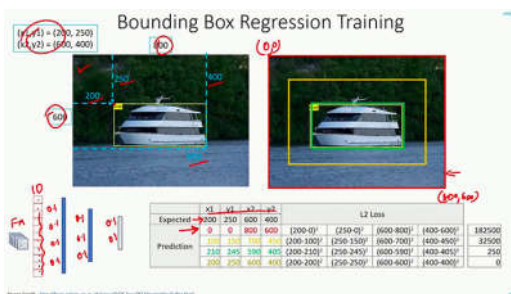


图3：目标检测中边界框回归算法

从根本上讲，我们可以将图像分类视为预测类标签。这种类型的模型不能转化为物体检测。对于我们来说，不可能为输入图像中 (x, y) 坐标边界框的每种可能组合构造一个类标签。我们需要依赖于另一种称为回归的机器学习模型。与产生标签的分类不同，回归使我们能够预测连续值。

在网络的顶部，放置一个具有四个神经元的完全连接层，分别对应于左上角和右下角 (x, y) 坐标。

给定四神经元层，请实现S型激活函数，以使输出返回的范围为[0, 1]。

在由 (1) 输入图像和 (2) 图像中对象的边界框组成的训练数据上使用均方误差或均值绝对误差等损失函数训练模型。

训练后，我们可以将输入图像呈现给边界框回归网络。然后，我们的网络将执行前向传递，然后实际预测对象的输出边界框坐标。

声纹识别网络

2.3 ECAPA-TDNN模型网络结构

ecapa_tdnm由conv1D+BN、SE-Res2Block、ASP+BN、FC+BN以及AAM-softmax等模块构成。

我们的模型基于ResNet架构的ResNet34实现，在特征相加时维度相等的情况下，一个残差块被定义为：

$$y = F(x, \{W_i\}) + x$$

其中 x 和 y 是层的输入和输出的向量。函数 $F(x, \{W_i\})$ 表示要网络模型学习的残差映射函数。

该网络的卷积层使用二维特征作为输入参数,并使用2D的CNN对其进行处理,和 x -vector向量拓扑的原理类似,在池化层中收集平均值和标准差统计信息。总之,不论网络的深浅,使用残差学习的好处:网络层数较浅时能够加快收敛,网络层数很深时也可以解决退化问题,使求解器能够找到更好的解。

2.4 声音及震动数据的预处理

声音的处理需要把采集的声音数据转化成特征图谱,并且需要做必要的预处理,比如:声音增强处理,降噪处理和对长音频做切片处理等。

我们采用了librosa语音信号处理的第三方库,Librosa是一个用于音频、音乐分析、处理的python工具包,一些常见的时频处理、特征提取、绘制声音图形等功能。

项目中我们使用了对声音文件的读取,对声音文件重采样等功能,包括处理采样率(sr)、帧移(hop_length)、频谱图或叫做语谱图(spectrogram)等操作。

2.5 声纹相似度识别算法

采集声音,经过预处理,通过ecapa_tdn就可以生成一个192维的特征值,我们把这个特征值作为声纹处理。下一步我们需要对不同声音输出的特征值做比对,判断声音的相似度。我们使用余弦相似度来计算声纹的相似度。余弦距离是一个近似意义上的距离度量公式,形容两个特征向量之间的关系还是有很大用处的。比如人脸识别,声纹比对等。

我们把正常设备声纹保存到声纹库中,新采集的声纹和声纹库通过余弦相似度计算公式计算新的声纹和声纹库中的声纹相似度,取声纹最接近的声纹作为分类,通常设置一个阈值,如果低于这个阈值,我们就认为不是这个声纹库的声音,作为异常声音处理。这样从样本的角度讲,只需要采集正样本和背景样本就可以了,在工程上这种声纹特征值的相似度的计算方法可以解决故障检测模型中负样本缺失的大难题,有非常积极的意义。

3 目标效果

实现智能化管控,智能运算,视频自动分析,对各类生产巡检内容的监控有效实现智能化监控管理。灵活管控、及时阻断实现PC端监控室、手机APP联动管理,自动警报,协助运营、实时监控,减少安全隐患。

3.1 保障人员安全 疫情防控、安全隐患排查(AI识别)

3.2 提高生产效率 工艺流程监测(AI识别)

3.3 维护财产安全 设备故障预测(拾音设备+声纹识别)

3.4 逐级预警提醒 分级分权、各负其责

3.5 一键定期体检 运筹帷幄,掌控厂区健康状况

3.6 AI+数据建模 洪湖水厂大数据模型,预见未知
传统巡检:制定“天/周、特殊天气”的不同巡检计划,约4小时巡检1次,采用AI智慧巡检实现7*24小时巡检。

致谢

感谢分公司领导对项目的全力支持,感谢洪湖水水质净厂运营团队提供大量的素材和业务场景,并深度参与项目的研发。感谢深圳水务科技提供的技术支持和指导。感谢编辑工作人员及审稿专家给予的大力帮助。

参考文献

- [1]基于声纹识别技术的调度下令验证系统研究[J]. 杨子,王吉文,王波. 电子器件. 2021,第6期
- [2]基于云计算的水务大数据平台系统设计与实践[J]. 罗贤伟,庞子山,谭松柏. 给水排水. 2022,第1期
- [3]基于注意力机制的卷积神经网络遮挡目标检测算法[C]. YANG Junzhi,杨钧智,WU Jinliang. 第七届高分辨率对地观测学术年会. 2020.
- [4]Research on the Application of Web Inspection System Based on CORS in City Inspection [C]. Xu Ding, Jiming Guo, Mingduan Zhou, CSNC 2012;China satellite navigation conference. 2012
- [5]D.Snyder,D.Garcia-Romero,G.Sell,D.Povey,andS. Khudanpur,“X-vectors:RobustDNNembeddingsforspeakerrecognition,” inProc.ICASSP,2018.