

基于多传感器融合的空气质量检测

李景富, 崔英杰

(黄淮学院, 河南 驻马店 463000)

摘要: 针对当前空气质量检测错误率高,检测效率低的难题,以获得理想的空气质量检测结果为目标,设计了基于多传感器融合的空气质量检测模型。首先,分析了当前空气质量检测的研究现状,找到引起空气质量检测效果差的原因;然后,采用多个传感器同时对空气质量检测数据进行采集,建立空气质量检测的学习样本;最后,采用机器学习算法对空气质量检测学习样本进行训练,构建空气质量检测模型,并采用验证数据与其他空气质量检测模型进行对比实验,该设计模型的空气质量检测精度超过95%,可以准确描述空气质量的变化特点,而且空气质量检测时间也远远少于对比模型,提升了空气质量检测效率,具有更加广泛的应用范围。

关键词: 空气质量; BP神经网络; 检测模型; 传感器阵列; 学习样本; 机器学习算法

中图分类号: TN911.1-34; TP391

文献标识码: A

文章编号: 1004-373X(2020)19-0171-04

Air quality detection based on multi-sensor fusion

LI Jingfu, CUI Yingjie

(Huanghuai University, Zhumadian 463000, China)

Abstract: In view of the high error rate and low detection efficiency in the air quality detection, an air quality detection model based on multi-sensor fusion is designed to obtain ideal air quality detection results. The current research status of air quality detection is analyzed, and the reasons of poor air quality detection effect are found. And then, several sensors are used to collect air quality detection data at the same time, so as to establish a learning sample of air quality detection. The machine learning algorithm is used to train the established learning sample to construct the air quality detection model. The verification data is compared with the data of other air quality detection models. The results show that the air quality detection accuracy of the proposed model is much higher than 95%, which can accurately describe the characteristics of air quality change. Moreover, its air quality detection duration is far shorter than that of the contrastive models, which improves the efficiency of air quality detection. Therefore, the proposed model has a wider application range.

Keywords: air quality; BP neural network; detection model; sensor array; learning sample; machine learning algorithm

0 引言

近年来空气污染天数有所增加,污染物主要是粉尘及其他微小颗粒物,不但直接影响人们生活质量,也使一些大型数据中心的通信设备故障率明显上升,通过分析,大部分故障与数据中心运行环境中灰尘超标有关。因此精度空气质量检测模型构建具有十分重要的研究价值^[1-3]。

当前主要采用传感器对空气质量数据进行在线采集^[4-6]。由于空气成分比较复杂,单一传感器采集的空气质量数据不完整,只能描述空气质量的片段、部分变化

特点,影响了后续空气质量的检测结果^[7]。采集空气质量数据后,同时需要引入一定的技术,建立空气质量检测模型。当前空气质量检测模型可以划分为两大类:线性检测模型和非线性检测模型^[8-10]。线性检测主要包括:聚类分析的空气质量检测模型、多元线性回归的空气质量检测模型,它们认为空气质量呈现一种固定的变化特点,包括线性、随机性、周期性、非平稳性变化特点,因此线性空气质量检测模型的缺陷相当明显,空气质量检测误差大,无法应用于实际中^[11-12]。在实际应用中,RBF神经网络的收敛性能不稳定,使得空气质量检测结果稳定性不够;最小二乘支持向量机的学习时间长,会对空气质量检测效率产生不利影响,因此空气质量检测有待进一步研究。

为了提高空气质量检测精度,本文设计了基于多传

收稿日期:2020-02-27

修回日期:2020-03-23

基金项目:河南省重点科技攻关项目(192102210285);驻马店市重点科技攻关项目(17230)

传感器融合的空气质量检测模型。采用多个传感器采集空气质量数据,利用BP神经网络建立空气质量检测模型,测试结果表明,本文模型的空气质量检测精度得到了大幅度提升,能够有效降低空气质量检测误差,减少了空气质量检测时间,改善了空气质量检测效率,空气质量检测结果要明显优于当前其他检测模型,验证了本文模型的优越性。

1 基于多传感器融合的空气质量检测模型

1.1 多传感器采集空气质量数据

因为空气质量的影响因素多,单一传感器采集的信息不完整,基于组合优化原则,采用多个传感器对同一个地方的空气质量数据进行采集,使各个传感器数据之间可以进行互补,提高了空气质量数据的采集精度,方便后续的空气数据检测建模。多传感器采集空气质量数据的原理如图1所示。

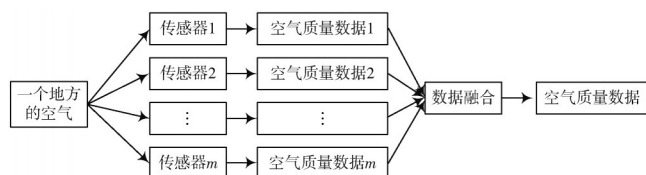


图1 多传感器采集空气质量数据的原理

1.2 BP神经网络

为了解决RBF神经网络和最小二乘支持向量机在空气质量检测过程中存在的缺陷,本文引入BP神经网络对空气质量检测进行建模。相对于RBF神经网络,BP神经网络的稳定性更高;相对于最小二乘支持向量机,BP神经网络的学习时间缩短,学习速度明显加快,是一种当前广泛应用的机器学习算法。对于一个空气质量检测问题,采用一个3层BP神经网络就可以对其变化特点进行高精度、快速逼近,BP神经网络结构可以采用图2进行描述。

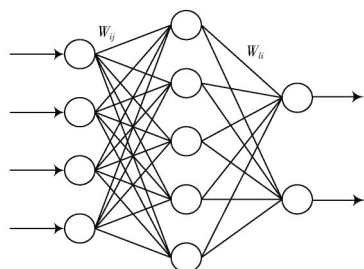


图2 BP神经网络的结构图

BP神经网络的输入数据可以表示为: $O_j^{(1)} = x(j)$, $j = 1, 2, \dots, n$,其中, n 表示空气质量的影响因素,隐含层的输入数据根据输入层的数据确定,而隐含层的输出数据又可以通过隐含层输入数据和映射函数 $f()$ 确定,隐

含层的输入数据和输出数据分别可以表示为:

$$\text{net}_i(k) = \sum_{j=0}^n W_{ij} O_j \quad (1)$$

$$O_i(k) = f(\text{net}_i(k)) \quad (2)$$

式中 W_{ij} 表示隐含层加权系数。

隐含层是BP神经网络最为关键的部分,其主要功能是对数据进行一定的变换,映射函数 $f()$ 具体定义为:

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3)$$

根据隐含层的输出数据确定输出层的输入数据,并且可以根据输出层的输入数据和映射函数 $g()$ 确定输出层的输出数据,输出层的输入数据和输出数据分别可以表示为:

$$\text{net}_l(k) = \sum_{i=0}^q W_{li} O_i(k) \quad (4)$$

$$O_l(k) = g(\text{net}_l(k)) \quad (5)$$

式中: W_{li} 表示隐含层加权系数;映射函数 $g(\cdot)$ 具体定义如下:

$$g(x) = \frac{e^x}{e^x + e^{-x}} \quad (6)$$

BP神经网络空气质量检测模型的实际输出和空气质量检测期望输出之间的误差为 $E(k)$,BP神经网络 $E(k)$ 对加权系数进行调整和修正,引入一个惯性项,可以得到加权系数调整形式为:

$$\Delta W_{li}(k) = -\eta \frac{\partial E(k)}{\partial \Delta W_{li}} + \alpha \Delta W_{li}(k-1) \quad (7)$$

式中: k 表示学习代数; η 为学习速率; α 为惯性系数,且有:

$$\frac{\partial E(k)}{\partial \Delta W_{li}} = \frac{\partial E(k)}{\partial y(k)} \cdot \frac{\partial y(k)}{\partial u(k)} \cdot \frac{\partial u(k)}{\partial O_i(k)} \cdot \frac{\partial O_i(k)}{\partial \text{net}_l(k)} \cdot \frac{\partial \text{net}_l(k)}{\partial W_{li}(k)} \quad (8)$$

式中 $y(k)$ 表示传递函数,并且存在:

$$\frac{\partial \text{net}_l(k)}{\partial W_{li}(k)} = O_i(k) \quad (9)$$

BP神经网络输出层的加权系数学习过程为:

$$\Delta W_{li}(k) = \alpha \Delta W_{li}(k-1) + \eta \delta_l O_i(k) \quad (10)$$

式中, δ_l 表示输出层阈值,具体可以表示为:

$$\delta_l = E(k) \frac{\partial y(k)}{\partial u(k)} \frac{\partial u(k)}{\partial O_l(k)} g'(\text{net}_l(k)) \quad (11)$$

式中 $g'() = g(x)(1 - g(x))$ 。

隐含层加权系数的学习过程为:

$$\Delta W_{ij}(k) = \alpha \Delta W_{ij}(k-1) + \eta \delta_i O_j^{(1)}(k) \quad (12)$$

式中, δ_i 表示隐含层的层阈值,可以表示为:

$$\delta_i = f'(\text{net}_i(k)) \sum_{l=1}^n \delta_l W_{li}(k) \quad (13)$$

式中 $f'(x)=(1-f^2(x))/2$ 。

1.3 基于多传感器融合的空气质量检测步骤

- 1) 采用多个传感器对空气质量数据进行采集,并去除每一个传感器数据中的无用信息。
 - 2) 将多个传感器采集的空气质量数据进行融合,同时对空气质量数据中的重复信息进行去除,减少空气质量数据的存储空间,获得高质量空气质量检测的样本数据。
 - 3) 将空气质量检测数据划分为学习数据和验证数据。
 - 4) 根据空气质量检测学习数据,对 BP 神经网络结构进行初始化。
 - 5) 采用 BP 神经网络对空气质量检测学习数据进行训练,建立空气质量检测模型。
 - 6) 采用建立的空气质量检测模型对验证数据进行检测,统计空气质量检测精度,并输出相应的结果。
- 综合上述可以得到基于多传感器融合的空气质量检测流程如图 3 所示。

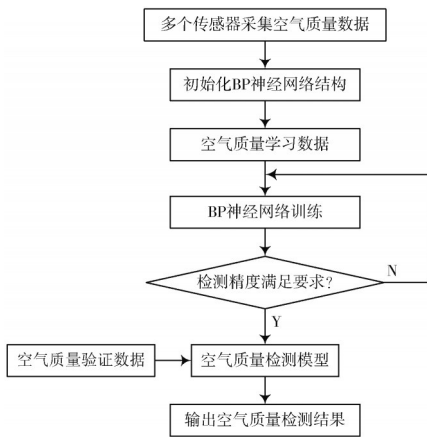


图 3 基于多传感器融合的空气质量检测流程

2 多传感器融合的空气质量检测的性能测试

2.1 测试环境

为了测试基于多传感器融合的空气质量检测模型的性能,采用 VC++ 6.0 编程实现多传感器融合的空气质量检测程序,采用的测试环境如表 1 所示。传感器数量为 10 个。

2.2 样本数据以及对比模型

为了使基于多传感器融合的空气质量检测结果具有可比性,选择单一传感器的空气质量检测模型进行对比测试,其建模工具分别为 RBF 神经网络和最小二乘支持向量机。因此有两个对比模型:RBF 神经网络和最小二乘支持向量机,为了使实验更加公平,可信度更高,

每一种模型进行 5 次仿真实验,每一次仿真实验的学习数据和验证数据的数量如表 2 所示。

表 1 空气质量检测的测试环境

环境参数	配置
CPU	4 核 3.0 GHz
RAM	16 GB
OS	WIN10
HD	1 000 GB

表 2 空气质量检测学习数据和验证数据的数量

实验编号	学习数据 /个	验证数据 /个
1	100	50
2	200	100
3	150	75
4	100	50
5	150	75

2.3 空气质量检测精度对比

统计 3 种空气质量检测模型的检测精度,结果如图 4 所示。从图 4 可以看出:RBF 神经网络的空气质量检测精度平均值为 86.25%,最小二乘支持向量机的空气质量检测精度平均值为 89.00%,本文模型的空气质量检测精度平均值为 96.04%。相对于对比模型,本文模型的空气质量检测精度平均值分别提高了 9.79% 和 7.04%,有效降低了空气质量检测误差。这是因为采用多个传感器对空气质量数据同时进行采集,采集的数据更加全面,并且引入 BP 神经网络对空气质量检测进行建模,获得了理想的空气质量检测结果,验证了本文空气质量检测模型的优越性。

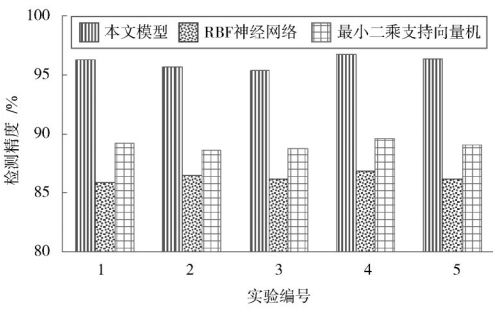


图 4 空气质量检测精度对比

2.4 空气质量检测效率对比

统计 3 种空气质量检测模型的检测时间,结果如图 5 所示。从图 5 可以看出:RBF 神经网络的空气质量检测时间平均值为 4.03 ms,最小二乘支持向量机的空气质量检测时间平均值为 6.19 ms,本文模型的空气质量检测时间平均值为 3.16 ms。相对于对比模型,本文模型的空气质量检测时间平均值分别减少了 0.87 ms 和 3.03 ms,加快了空气质量检测速度,这是因为 BP 神经

网络的学习速度更快,提高了空气质量的检测效率。

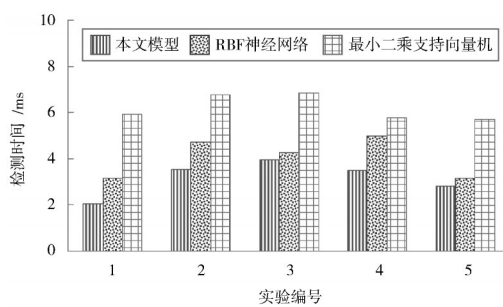


图5 空气质量检测时间对比

3 结 语

为了解决当前空气质量检测错误率高,检测效率低的难题,本文提出了基于多传感器融合的空气检测模型。首先采用多个传感器同时对空气质量数据进行采集,建立空气质量检测学习样本;然后采用机器学习算法对空气质量检测学习样本进行训练,构建空气质量检测模型;最后在相同环境下与其他空气质量检测模型进行了对比实验,本文模型的空气质量检测精度超过95%,可以准确描述空气质量变化特点,而且空气质量检测时间也少于对比模型,提高了空气质量检测效率,为空气质量检测提供了一种建模工具,对数据中心智能防尘具有重要参考价值。

作者简介:李景富(1981—),男,河南确山人,硕士,高级实验师,主要研究方向为信息安全、协同过滤、嵌入式系统。

崔英杰(1992—),男,河南驻马店人,硕士,助教,主要研究方向为物联网、信息技术。

(上接第170页)

- [2] 朱光,丰米宁,张薇薇.激励机制下图书馆信息安全管理投入意愿研究:基于演化博弈的视角[J].数据分析与知识发现,2018(6):13-24.
- [3] 田华,鄢喜爱,杨金民.图书馆云存储系统中一种自适应容错策略[J].计算机工程与应用,2019,55(7):95-99.
- [4] 王蕾.媒介融合语境下高校图书馆如何联合开展移动信息服务:评《大数据时代高校图书馆信息服务创新研究》[J].新闻爱好者,2018(1):115-116.
- [5] 孙战彪.SOLOMO环境下图书馆信息资源协同建设研究[J].现代情报,2017,37(12):110-116.
- [6] 谷参.基于分布式结构的图书馆信息检索服务系统研究[J].现代电子技术,2017,40(1):83-85.
- [7] 叶小榕,邵晴.结合物联网和室内定位的手机图书馆推荐系统[J].科技导报,2016,34(23):127-136.
- [8] 张静端.基于GIS技术的高校图书馆座位管理系统[J].东华大

学学报(自然科学版),2016,42(2):242-247.

- [9] 许磊,夏翠娟.第三代图书馆服务平台的元数据管理:以FO-LIO的Codex方案为例[J].中国图书馆学报,2020,46(1):99-113.
- [10] 谭龙江.基于读者活动的个性化图书馆推荐系统[J].系统工程,2015,33(8):154-158.
- [11] 张凤霞.大数据环境下基于云计算的图书馆用户信息挖掘技术研究[J].现代电子技术,2019,42(21):36-40.
- [12] 陆韩.面向OA资源的图书馆知识发现系统应用研究[J].图书馆工作与研究,2019(10):83-88.
- [13] 侯欢欢,杨志萍,侯雪婷.可行能力视角下图书馆信息素养扶贫路径研究[J].图书馆学研究,2019(18):69-74.
- [14] 王栋,袁伟,吴迪.基于WiFi物联网的图书馆环境监测系统[J].计算机科学,2018,45(z2):532-534.
- [15] 李金.基于云计算的数字图书馆智能问答系统设计[J].制造业自动化,2019,41(8):136-140.

作者简介:吴云(1977—),女,图书馆馆员,主要从事图书采编管理、信息服务研究。