

基于 WDNn 的温室多特征数据融合方法研究

孙耀杰¹ 蔡 昱^{1,2} 张 馨³ 薛绪掌³ 郑文刚² 乔晓军²

(1. 河北工业大学电子信息工程学院, 天津 300401; 2. 北京农业智能装备技术研究中心, 北京 100097;
3. 北京农业信息技术研究中心, 北京 100097)

摘要: 目前物联网监测产品在温室生产中大量应用产生海量数据,但现有用于温室数据融合算法对高维特征及混合特征(数据同时包含稀疏特征和连续特征)处理精度较低且泛化能力较弱,鲜有利用深度学习模型对温室数据进行顶层融合并提供准确的决策信息。本文提出了一种基于宽-深神经网络(Wide-deep neural network, WDNn)的两级温室环境数据融合算法。首先利用温室内多点多特征数据训练 WDNn 深度学习模型,输出形式为多点单特征,再将该输出数据按照少数服从多数原则进行融合,得到温室环境状态的整体评估结果。试验结果表明,该融合方法对预测集中混合特征的决策准确率达到 98.90%,融合特征类型的增加,可用于监测参数更多、环境更复杂的温室,将 WDNn 模型用于温室混合数据融合是可行有效的,在保证决策精度的同时丰富了可融合特征类别,进一步提升温室融合系统的智能化程度,对温室智能调控提供有效技术支撑。

关键词: 温室; 数据融合; 无线传感网络; 深度学习; 宽-深神经网络

中图分类号: S625.5⁺1; TP391.4 **文献标识码:** A **文章编号:** 1000-1298(2019)02-0273-08

Multi-feature Data Fusion Method of Greenhouse Based on WDNn

SUN Yaojie¹ CAI Yu^{1,2} ZHANG Xin³ XUE Xuzhang³ ZHENG Wen'gang² QIAO Xiaojun²

(1. School of Electronic Information Engineering, Hebei University of Technology, Tianjin 300401, China
2. Beijing Research Center of Intelligent Equipment for Agriculture, Beijing 100097, China
3. Beijing Research Center for Information Technology in Agriculture, Beijing 100097, China)

Abstract: The IoT monitoring products are widely used in greenhouse production, which could generate massive data. The existing data fusion algorithms for greenhouses had low fusion accuracy and weak generalization capability for high-dimensional features and mixed features (combined with sparse features and continuous features). It was rare to use the deep learning model to top-level fusion of greenhouse data and provide accurate decision information. Aiming at the above problems, a two-level greenhouse environment data fusion algorithm was proposed based on wide-deep neural network (WDNN). Firstly, integrating multi-point multi-features mixed data in the greenhouse and marking the data categories. Then the constructed training set and test set were input into the WDNn deep learning model for 2000-step iteration training. The model structure was set as 7-100-50-7, the output form was multi-point single feature, which was the first-level fusion result as decision information of each area of the greenhouse, and then the output data was second-level fusion according to the minority obeyed majority principle, and the overall evaluation decision of the greenhouse environmental state was obtained. For comparison purposes, the other three fusion models were trained as deep neutral network (DNN), BP neural network (BPNN) and random forest (RF). The experimental results showed that the loss value of the initial segment of the WDNn network was higher than that of DNN network, but the loss function curve had a faster rate of decline and the model parameters were better. The accuracy of the model after training was 4.32 percentage points higher than that of DNN, but the training time was increased by 21.36%; the accuracy of BPNN model was 82% and its parameter optimization was the slowest, parameter optimization required more iteration steps; RF model training speed was the fastest, but its model fusion accuracy was 3.39

收稿日期: 2018-09-05 修回日期: 2018-10-27

基金项目: 国家重点研发计划项目(2017YFD201503)和北京市农林科学院科技创新能力建设专项(KJCX20170204)

作者简介: 孙耀杰(1960—),男,教授,博士,主要从事通信与测控技术研究,E-mail: sun678@126.com

通信作者: 张馨(1981—),男,副研究员,博士,主要从事设施环境智能测控研究,E-mail: zhangx@nrcita.org.cn

percentage points lower than that of WDNN. The fusion accuracy was insufficient; above comparison results proved that it was feasible and excellent to use the WDNN model to fuse the mixed data in the greenhouse. Inputting the mixed situation information contained the sensor anomaly and meteorological data under various conditions into the fusion system, then the context decision rate reached 98.90%. The realization of the WDNN fusion system could be used to monitor greenhouses with more parameters and more complex environments, and enrich the fusion feature categories while ensuring the accuracy of decision-making. It could further improve the intelligence degree of the greenhouse fusion system.

Key words: greenhouse; data fusion; wireless sensor networks; deep learning; wide-deep neural network

0 引言

随着设施园艺面积的逐年增加,为生产环境调控提供智能决策成为当前人工智能在温室生产中研究应用的核心问题之一。若对数据未进行融合而选择直接判断提供决策信息,仅在少量输入特征时可行,随着输入特征的增加,直接判断法的逻辑复杂度随着特征的增加呈指数增长且无学习能力,降低了决策系统的可发展性和可扩展性,造成日后维护和升级成本大大增加,因此高效的数据融合算法显得格外重要。目前国内温室数据融合算法较为单一,鲜有将深度学习模型与温室数据融合相结合。深度学习以强大的数据分类及拟合优势,能够为温室复杂多变环境数据实时精确融合提供新的算法,掌握温室环境变化规律,对温室作物的生长和温室环境的调控具有重要意义。

当前温室数据融合研究主要分为数据级、特征级和决策级3个层次。其中数据级数据融合^[1]在温室环境监测系统广泛应用,对获取的原始数据进行加权平均比传统平均算法数据精度提高7%^[2]。进一步优化算法可将指数平滑和新型幂均方融合算法结合,使实际运行时间缩短83.6%的同时,数据方差比算术平均降低0.027^[3]。在面对数据异常情况时,将分布图算法与卡尔曼滤波相结合,增加系统对缺失数据处理的能力且处理精度进一步提高^[4]。数据级融合适合硬件设备进行传输和处理,但融合后数据量庞大,从而产生了特征级数据融合^[5-6]。因温室环境复杂导致监测数据常出现异常点、冗余或传输干扰等现象,多数特征级融合与数据级融合算法相结合,如卡尔曼滤波器与贝叶斯估计相结合^[7],可改善数据融合前的一致性以及利用贝叶斯算法融合后的不确定性。或将自适应加权平均和D-S算法相匹配^[8],改进后的D-S算法时间复杂度由传统的 $O(a^n)$ 下降到 $O(a^2n)$ ^[9],在减少数据冗余的同时,提高了系统的运行效率,但存在模型参数难以确定等问题。决策级温室数据融合集中在专家系统和BP神经网络^[10-11]的构建,专家系统利用推

理机对已有条件做出精确判断及相关决策^[12-13]。神经网络可利用ANN模型对温室温度及湿度进行融合,融合结果与真值对比,均方误差(MSE)分别达到0.04℃和0.075^[14]。决策级数据融合层次位于顶层,具有高容错性和强鲁棒性的特点,但目前主流的决策级数据融合算法存在处理数据维度有限,收敛速度慢且泛化能力差等问题。

针对用于温室融合算法老旧、智能化程度低、融合数据种类单一及模型融合精度不足等问题,本文提出基于WDNN网络的两级温室数据融合体系,利用深度学习模型的强大非线性拟合及泛化能力,通过温室传感器网络采集数据并训练模型,以探索对多点多维混合特征的决策级融合,给出温室各区域及整体环境的决策信息。

1 温室数据获取与深度融合结构设计

试验于2017年11月20日18时在北京市农林科学院玻璃连栋温室内架设仪器采集数据,试验场地示意图如图1所示。

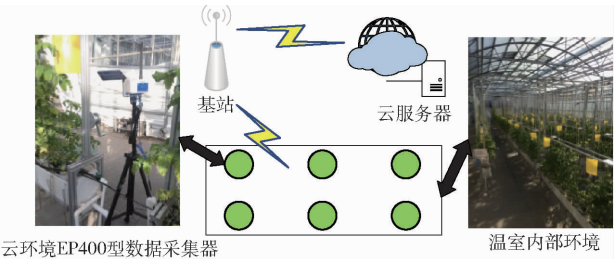


图1 温室环境监测网络布局与架构图
Fig.1 Greenhouse environment monitoring network layout and architecture diagram

该温室单跨南北长度28.5 m,东西宽度10.5 m,总高6 m,肩高4 m左右,温室内种植作物为番茄^[15-16],为符合番茄种植要求,温室采用相应调控措施,其中最大湿度阈值设定为80%,当湿度超过阈值会开启风扇进行排湿,但环境调控效果存在滞后现象,故湿度不会立刻降到80%以下,需要一段时间缓冲,因此本试验所获取湿度最大值会略高于80%,但不会使其达到100%,可有效避免番茄产生病害。温室环境监测设备采用国家农业智能装备技

术中心研发的温室云环境 EP400 型数据采集器,每个数据采集器可测量空气温度(量程: $-40 \sim 80^{\circ}\text{C}$,精度 $\pm 0.5^{\circ}\text{C}$)、空气相对湿度(量程: $0 \sim 100\%$,精度: $\pm 3\%$)、土壤温度(量程: $-40 \sim 80^{\circ}\text{C}$,精度: $\pm 0.5^{\circ}\text{C}$)、土壤湿度(量程: $0 \sim \text{饱和}$,精度: $\pm 3\%$)、 CO_2 浓度(量程: $0 \sim 2 \times 10^{-3}$,精度: $\pm 5 \times 10^{-5}$)及光照强度(量程: $0 \sim 100 \text{ klx}$,精度: $\pm 10 \text{ klx}$),以 30 min 时间间隔通过 GPRS 网络发送到远端云服务器,数据采集器安装在三脚架上,距离地面高度为 1.3 m,各数据采集器纵向间距为 5 m,横向间距为 4.8 m,

对称均匀地放置在温室中,如图 1 中绿点所示。设置数据采集间隔为 30 min 一次,既保证了温室作物环境的有效监测、减少硬件传输损耗,又可保证模型训练的较大数据需求量,且无需在底层进行数据级的数据融合,降低系统复杂度,使数据的信息量没有损失,也保留了温室作物生长环境的原始信息。

采用 WDNN 深度学习模型对连栋温室内多点多特征数据进行决策级的数据融合。构建了两级温室数据融合系统,如图 2 所示。

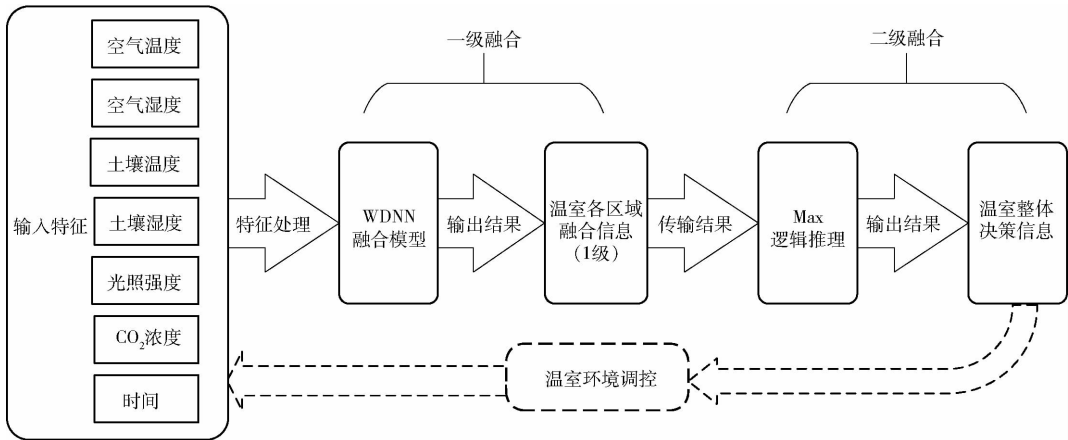


图 2 两级温室数据融合模型

Fig. 2 Two-level greenhouse data fusion model

图 2 中,输入特征为温室内部环境监测数据,本文未考虑外部气象数据等间接特征(即调控设施或外部天气变化时会引起温室内部环境状态改变),在可直接获取温室内部各环境变量的前提下,已经包含了其他因素对温室内部环境的影响,若添加间接特征会增加数据冗余度,降低所训练模型的泛化能力。一级融合为该系统的核心处理层次即第 1 次数据融合,利用深度学习模型对 6 个数据采集传感器传输的多源异构数据进行处理,并分别给出各个区域作物的生长环境状态。二级融合处理将一级处理结果进一步融合,该融合体系将二级融合称为 Max 逻辑推理,按照少数服从多数原则进行二次融合,当区域融合结果相同数量占据总量 1/2 及以上时,该区域融合结果代表温室整体融合结果(若胁迫结果与正常结果相等时,取胁迫结果,此设定可减少作物受到伤害)。第 2 次数据融合后,给出温室整体的环境状态并提供温室环境调控建议,以达到对作物生长环境进行调控的要求。图 2 虚线模块为温室硬件调控模块,主要负责接收根据综合决策信息而提供的调控建议,并实施相应调控策略,以虚线画出表示本文研究不涉及硬件层面,但却与硬件调控紧密结合。二级融合架构可对无线传感网络数据进行分层融合,且多级融合符合模块化设计理念,可

对架构作出相应调整以更好地适用于更复杂的融合环境,提高了融合系统的泛化能力。

2 传感器网络数据预处理

目前温室数据融合的预处理主要集中在底层,会采用较小的数据采集间隔(多为 1 min 间隔)以提高数据量,但考虑到温室无线传感网络的功耗和传输效率,大多预处理会剔除或滤掉异常数据^[17-21],减少了数据本身的信息量,信息量的减少会造成对所监测对象的决策和分析产生较大误差。在利用 WDNN 网络模型进行顶层数据融合时,为了保留原始数据的全部信息,本试验并未对数据进行平滑处理(平滑或其他预处理会损失特征的信息量),而是对获取的全部原始数据进行标记,且在此基础上增加了时间特征,丰富的数据信息可增加融合后的决策准确度。

训练模型首先需构建训练集和测试集,因两种数据集构建方式一样故以训练集为代表进行描述。首先获取原始数据,为传感器实时传送并记录到数据库的环境监测数据,因融合过程是对输入数据进行分类,故训练集中每条数据都需进行类别标记。数据与标签的合理匹配需明确各输入特征间的权重关系,进行多特征分析,找出主导特征。各参数走势曲线如图 3 所示。

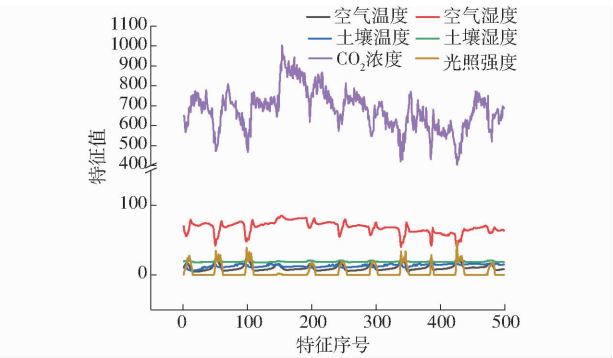


图3 温室环境特征

Fig. 3 Features of greenhouse environment

由图3可以看出,温室各环境特征间具有很强的规律性和周期性,特征之间具有明显的正相关或负相关关系,如温度、湿度两个参数具有强负相关特性,温度的增长会导致湿度的下降,温度、湿度、CO₂浓度和光照强度均为植物生长的重要环境参数,每个特征都应在融合系统中分配相应的权重,由图3可知,温度与其他监测特征具有强耦合关系且可控性高,所以以温度特征为基准去分析,可以更准确地反映其他参数的分布及变化趋势。

其次需增加训练集的数据量,将异常数据加到训练集中,可使模型对异常值进行判别,为全面准确地描述温室内各区域及整体的环境参数分布,需将6个数据采集点数据全部合到一起,可使训练集包

含整个温室内部的数据,处理后训练集部分数据如表1所示。表1中第1列时间变量是稀疏特征(特征值为1、2、3、4和5)。其中1、2、3分别代表上午、中午和下午,对应:06:00—10:00、10:00—14:00、14:00—18:00,只用来描述日间温室环境正常及胁迫数据;4代表日间,对应:06:00—18:00,只用来描述日间传感器异常数据;5代表夜间,对应18:00—06:00,用来描述夜间温室环境正常、胁迫及异常数据,用于细化不同时间段作物对环境的需求;第2~7列变量为数值数据,为各气象参数;第8列为分类类别,根据番茄适宜生长环境需求及具体监测数据情况划分为:日间温度过低、湿度过高、CO₂浓度较高、光照强度过低;日间温度较低、湿度较高、CO₂浓度正常、光照强度较低;日间各环境参数正常;日间温度较高、湿度较低、CO₂浓度较低、光照强度较高;夜间温度较低、湿度较高、CO₂浓度较高、无光照;夜间各环境参数正常;传感器网络数据异常共7个类别,以上类别涵盖了24 h各情况下的温室监测数据,适合WDNN网络对数据进行深度融合;此外,表1中1~4行为白天4个类别的典型数据示例;5、6行为夜间2个类别的典型数据示例;7~9行为传感器异常数据的典型示例(日间异常数据不进行细分,用4统一进行描述)。本文所用训练集、测试集及预测集均为实时数值。

表1 训练集数据示例

Tab. 1 Data example of training set

时间	空气温度/℃	空气湿度/%	土壤温度/℃	土壤湿度/%	CO ₂ 浓度	光照强度/klx	分类类别
1	7.72	79.02	6.25	19.85	9.67×10^{-4}	17.6	0
2	14.14	63.30	7.12	19.85	9.87×10^{-4}	24.1	1
3	18.46	56.23	12.68	22.47	6.53×10^{-4}	21.8	2
2	23.73	49.77	14.18	28.30	7.04×10^{-4}	32.6	3
5	7.07	71.87	12.25	18.31	1.01×10^{-3}	0	4
5	10.59	75.57	15.87	19.02	9.02×10^{-4}	0	5
4/5	100.00	100.00	-0.06	0.00	9.85×10^{-4}	0	6
4/5	18.75	12.62	-0.06	0.00	8.42×10^{-4}	0	6
4/5	-2.06	23.72	-3.75	0.00	4.42×10^{-4}	0	6

处理后的训练集和测试集均为7个特征和7个分类类别的混合数据集,测试集需要严格按照训练集的规格进行整理,才能精确地判断所训练模型的性能,最后需将整理好的Excel格式训练集和测试集转换成csv格式,方便程序对数据的读取及处理。

3 深度学习融合模型构建

3.1 普通DNN分类模型

传统深度学习网络^[22](Deep neural networks, DNN)即含多个隐含层(大于等于2层)的人工神经网络,理论上单个隐含层的多层感知机(Multilayer

perceptron, MLP)只要具有足够多的隐含层节点就可以拟合任何非线性函数,但是过多隐含节点极易导致优化算法无法确定模型最优参数或训练过拟合,无法进行泛化及应用。2006年HINTON等^[23]提出的深度学习算法很好地解决了上述问题,模型结构示意图如图4所示。

图4中的模型具有多隐含层结构,相比于传统的MLP模型而言具有更强的数据拟合能力,GOODFELLOW等^[24]试验表明,增加隐含层节点数量对模型的性能提升作用有限,模型深度的增加(增加隐含层层数)不仅可以减少每个隐含层节点

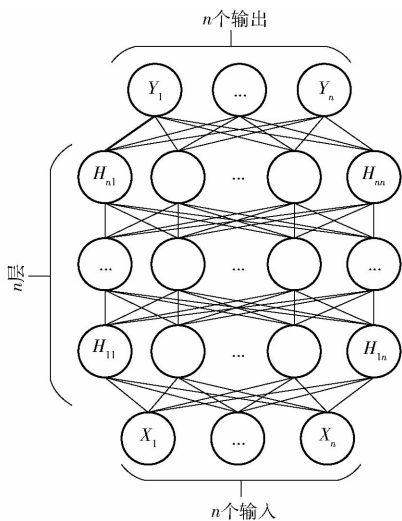


图 4 DNN 网络模型
Fig.4 DNN network model

的数量,还能增加对数据的处理和拟合能力。所以本文选择利用深度学习模型进行温室监测数据的融合和决策,可以很好地解决 MLP 隐含层节点过多导致收敛速度过慢及局部极小化的问题。

3.2 改进后的 WDNN 模型

DNN 深度学习模型可以较好地对多特征的数据进行分类,但是由于其模型结构限制和网络特性,使其不易处理稀疏特征,无法对特征之间的强耦合性进行关联记忆,当输入数据出现稀疏特征和关联特征时,不能快速并准确地对该关联特征进行参数优化。

Google 于 2016 年提出了宽-深神经网络 (Wide-deep neural network, WDNN)^[25]用于 Google 商店的 APP 推荐服务,该模型改进之处在于从原来 DNN 深度模型的基础之上联合训练了一个 wide 模型 (线性模型)用于共同对数据进行分类决策。wide 网络数学模型表示为

$$y = W^T X + b \tag{1}$$

其中 $X = (x_1, x_2, \dots, x_n)$ $W = (w_1, w_2, \dots, w_n)$

式中 X ——特征向量 W ——模型参数

b ——模型偏置

y ——wide 网络输出

wide 网络最重要的转换为交叉乘积变换,定义为

$$\phi_k(X) = \prod_{i=1}^d x_i^{c_{ki}} \tag{2}$$

式中 c_{ki} ——布尔变量,只有当第 i 个特征是 ϕ_k 第 k 次转换的一部分时为 1,其余条件下为 0

wide 模型善于处理稀疏特征,通过对特征进行交叉乘积来记忆特征之间的相关性且需要较少的模

型参数,记忆性可以理解为能够自主学习频繁共同出现在一起的特征并当该关系再次出现时可以发现并利用,善于处理特征之间的固定组合。

deep 网络具有更好的泛化能力,可以关联或发现之前几乎没有出现的特征组合,善于将隐藏的特征属性进行融合,减少了特征工程的复杂度。

deep 网络数学结构:

(1) 输入层为输入数据的各个特征。

(2) 每个隐含层节点均由激活函数构成,为了克服梯度消失等问题, DNN 模型构建时默认函数为整流线性单元 ReLu,函数定义为

$$g(z) = \max(0, z) \tag{3}$$

式中 z ——激活函数输入值

(3) 多分类模型输出层多为 SoftMax 函数,形式为

$$f(z_i) = \frac{\exp(z_i)}{\sum_j \exp(z_j)} \tag{4}$$

(4) DNN 模型的整体函数表达式为

$$f(x; W, c, w, b) = w^T \max(0, W^T x + c) + b \tag{5}$$

(5) 优化函数默认使用 Adagrad 算法

$$\theta_{i,t+1} = \theta_{i,t} - \frac{\eta}{\sqrt{G_{i,t}} + \varepsilon} \nabla_{\theta_{i,t}} J(\theta) \tag{6}$$

式中 ε ——一个极小值,防止分母为 0

η ——学习速率 θ ——模型系数

$\nabla_{\theta_{i,t}} J(\theta)$ ——模型第 i 个系数 θ 第 t 步的梯度

deep 网络的算法构建主要由以上函数组成,函数 1 为隐含层节点的激活函数,负责对特征进行非线性变化以达到进行非线性拟合的目的,其中层内节点无连接,层间节点全连接;函数 2 为深度学习模型常用的多分类函数,用于搭建模型输出层;函数 3 为特征经过一层隐含层后的数学表示,即前一层输出经过 ReLu 函数进行非线性变化后传递给下一层,模型深度的增加就是该表达式的多层嵌套;函数 4 为常用的梯度下降算法,具有动态调整学习速率的特点,能够快速优化模型参数。因此将以上两个模型进行合并可以达到在不改变深度网络对数据融合能力的同时加强 WDNN 模型对关联特征的识别和记忆,改进后的模型示意图如图 5 所示。

WDNN 模型的训练采用联合训练,即训练阶段将 wide 和 deep 网络的参数一起优化,再共同对未知数据进行融合,改进后数据融合的数学表达式为

$$P(Y|x) = \text{SoftMax}(W_{\text{wide}}^T [x, \phi(x)] + W_{\text{deep}}^T a^{(l_f)} + b) \tag{7}$$

式中 Y ——分类类别标签

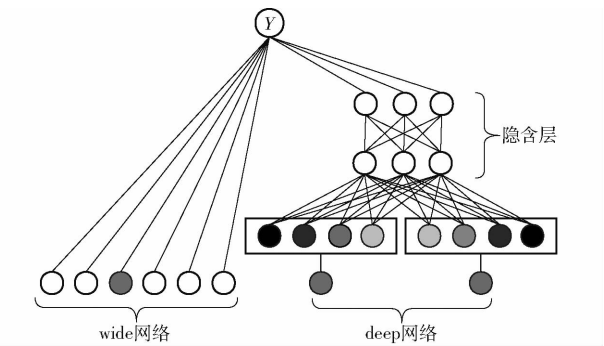


图5 WDN网络结构示意图

Fig.5 WDN network structure diagram

SoftMax(·)——多类别分类函数
 $\phi(x)$ ——原始特征 x 的交叉乘积转换
 W_{wide} ——wide 模型的权重向量
 W_{deep} ——deep 网络的权重向量

本文将 WDN 模型和温室作物环境数据相结合进行训练可以得到数据融合模型,其结构示意图如图 6 所示。

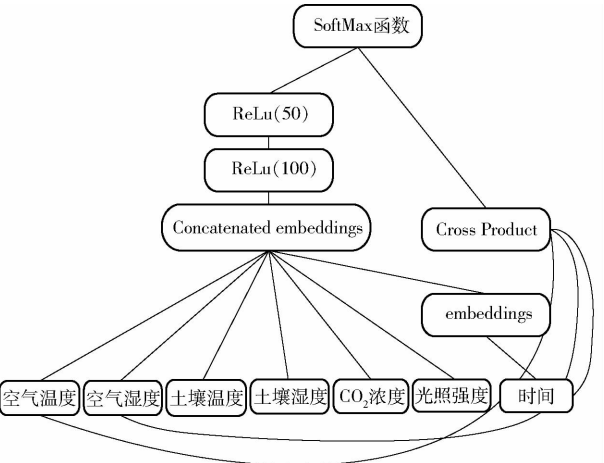


图6 WDN网络融合温室数据示意图

Fig.6 Schematic of WDN network fusion of greenhouse data

4 试验与分析

4.1 数据采集及准备

试验选取了 6 个数据采集点,从 2017 年 12 月 23 日至 2018 年 1 月 2 日的 4 858 条数据用于模型训练和测试,其中训练集 4 009 条数据,测试集 849 条

数据,每条数据均为 30 min 间隔、24 h 连续采集。将下载数据进行标记并将训练集和测试集用于模型的搭建。融合系统搭建环境为 MacOS High Sierra 操作系统(版本为 10.13.3),模型实现基于 Google 开源的 Tensorflow 工具包(版本为 Develop r1.1),编程语言为 Python(版本为 3.6),集成开发环境(IDE)为集成在 Anaconda 中的 jupyter notebook。

搭建模型前需对数据进行特征处理,按照番茄种植规范^[26]将温度划分为 6 个区间段,分别为日间:0~10℃、10~15℃、15~20℃、20~30℃,夜间:0~8℃、8~12℃;湿度划分为 3 个区间段(不区分日间和夜间):30%~60%、60%~70%、70%~80%;依据划分标准对训练集和测试集添加标签,然后与时间特征进行交叉乘积输入到 wide 模型;其余气象特征不进行区段划分;将实际值特征直接输入到 deep 模型中并与 wide 模型进行联合训练,特征对应分类类别为 7 类(数字 0~6),每一类别对应一个融合结果(各区域作物当前环境状态融合后的决策信息)及温室调控建议,成功构建用于温室数据融合的深度神经网络,用户可以根据融合信息实时掌握各区域作物的各项参数指标及温室整体的调控信息,实现温室作物环境数据融合的信息化和智能化。

4.2 模型构建及对比分析

为对比测试,分别训练了 DNN 和 WDN 两种深度学习模型及 BP 神经网络(BP neural network, BPNN)和随机森林(Random forest, RF)两种常用分类模型,选取 DNN 可实现在网络深度相同及对应算法一致的条件下进行对比测试,利用 BPNN 和 RF 可对比相同数据集条件下不同分支模型的融合精度情况,模型参数对比如表 2 所示。

由表 2 可以看出,由于 WDN 网络为混合模型,所以增加了一个 wide 网络的优化算法 FTRL,因具有良好的稀疏性及收敛特性,故适用于线性模型的优化训练,两个 deep 网络部分的模型参数配置均相同;BPNN 模型为 3 层结构,只含一个隐含层,层中部署 100 个隐含节点,为避免 ReLu 与 SoftMax 直接相连时训练过程出现权重为零值现象,激活函数

表 2 4 种模型参数对比

Tab.2 Comparison of four models parameters

模型	模型深度	节点数量	激活函数	优化算法	模型类别
DNN	2	100/50	ReLu	Adagrad	deep
WDNN	2	100/50	ReLu	Adagrad/FTRL	wide 和 deep
BPNN	1	100	Tanh	Gradient descent	MLP
RF	Auto	50	None	Cart	Ensemble learning

注: Cart (Classification and regression tree): 分类回归树; Ensemble learning: 集成学习; Gradient descent: 梯度下降; MLP (Multilayer perceptron): 多层感知机。

选择 Tanh 即可;RF 模型的决策树深度选择自由生长,对应算法选取 Cart,弱分类器个数为 50 个,设定为有放回随机挑选特征。

模型参数预设完成后进行训练,模型训练步骤均为 2 000 步,训练损失函数相同,对比如图 7 所示。损失函数定义为

$$H(y,a) = -\frac{1}{n} \sum_n y \lg a + (1-y) \lg (1-a) \quad (8)$$

式中 a ——预测变量

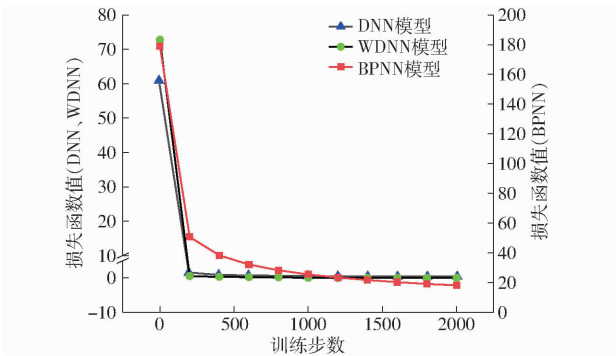


图7 DNN、WDNN 和 BPNN 模型损失对比

Fig.7 Loss comparison of DNN, WDNN and BPNN models

由图 7 可以看出,虽然 WDNN 网络在训练初段损失值高于 DNN 网络,但是曲线下落速率更快,模型参数优化效果更优,训练后的模型准确率更高;BPNN 模型训练损失值远高于其他 2 个模型,参数优化较慢。4 个模型性能对比评估如表 3 所示。

表 3 4 个模型性能评估

Tab.3 Performance evaluation of four models

模型	训练时间/s	测试准确率/%	损失函数值
DNN	12. 20	94. 58	0. 16
WDNN	14. 80	98. 90	0. 01
BPNN	15. 32	82. 00	18. 13
RF	3. 00	95. 51	

通过数据对比可知,WDNN 模型的训练时间相比于 DNN 网络增加了 21. 31%,这是由于 WDNN 网络需联合训练 wide 和 deep 网络,模型复杂度较高导致,但模型的测试准确率高了 4. 32 个百分点,且参数优化更快、损失函数值更低;BPNN 模型训练时间最长,融合精度最低,且训练所需步数较多;RF 模型虽训练时间最短,但模型预测准确率比 WDNN 降低了 3. 39 个百分点,融合精度不足;结果表明 WDNN 网络对于强耦合性的温室监测数据具有更好的拟合能力,使融合后数据的决策分析更加准确。

4.3 融合结果

将搭建完成的 WDNN 深度学习模型用于融合实时的温室监测数据,选取了 2018 年 1 月 500 个全新数据构建预测集,预测集包含传感器异常和区域

环境差异等各种条件下温室环境数据,用于测试温室数据融合的准确率,融合结果如图 8 所示。

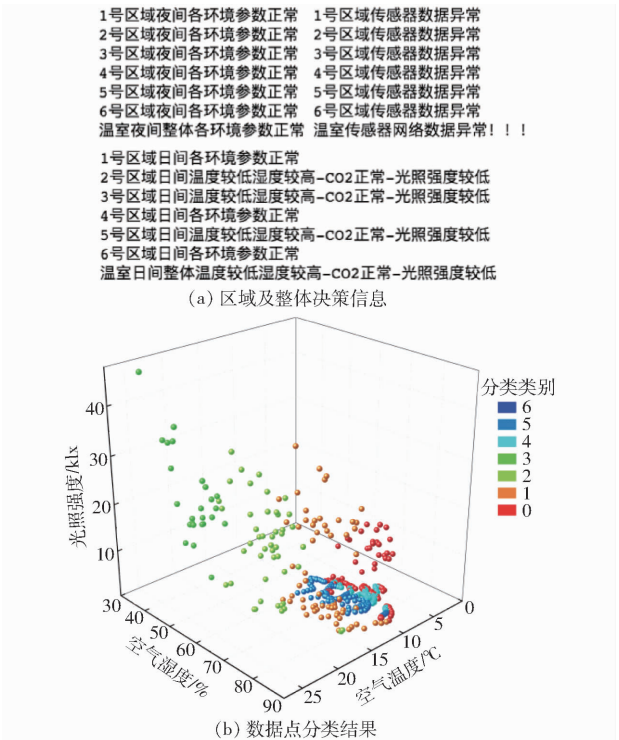


图 8 融合结果

Fig.8 Charts of fusion result

由图 8a 可以看出,该模型既可发现异常数据并进行警告,也可在精确融合各区域情景信息的同时,按照少数服从多数的原则对区域融合结果进行二级融合,得到温室整体环境综合决策(若胁迫结果与正常结果相等时,取胁迫结果,此设定可减少作物受到伤害);图 8b 为降维后的模型分类结果,可看出各簇数据分类效果明显,个别交叉数据点为降维后视觉误差所致;测试表明融合准确率高达 98. 00%,达到预测期望结果。

4.4 模型通用性验证

为验证 WDNN 融合模型的通用性,随机选取了 2018 年 2—6 月共计 200 个数据(随机选取数据为春季和夏季监测数据,可同时验证模型对季节的泛化能力)进行融合系统通用性验证。融合结果如表 4 所示,其中正确融合数据量为 190 个,对应融合

表 4 模型通用性验证

Tab.4 Verification of model generalization 个

月份	数据总量	正确融合数量	错误融合数量
2	40	39	1
3	40	39	1
4	40	37	3
5	40	38	2
6	40	37	3
总计	200	190	10

准确率为 95.00%, 融合精度相对于训练时有所下降, 主要原因为春夏气象特征分布与冬季差别较大, 会对模型融合精度产生一定影响, 但仍然高于 DNN 和 BPNN 模型且与 RF 模型几乎持平, 能够满足温室作物环境数据融合的需求。

5 结论

(1) 温室数据融合可以有效地监测并给出温室作物生长环境的决策信息。分析各深度学习模型的优缺点, 选取 WDNN 网络模型, 利用其对数据的强大非线性拟合及泛化能力, 根据无线传感网络获取

的多点监测数据建立了一个两级温室环境数据融合体系, 给出温室整体的作物环境决策信息。

(2) 在深度网络参数配置相同时, 改进后的 WDNN 网络比 DNN 和 BPNN 网络参数优化更快, 损失函数值更低; 改进后的模型输入特征不局限于连续实值数据, 对含有稀疏特征的混合信息依然可以进行精确融合, 且融合精度较 DNN 模型高 4.32 个百分点、高 BPNN 模型 16.90 个百分点、高 RF 模型 3.39 个百分点; WDNN 模型除了具有高精度的融合能力外, 还兼具良好的通用性, 可应对复杂多变的温室作物环境。

参 考 文 献

- [1] 李志刚, 刘丹丹, 张小栓. 基于分簇数据融合的农产品冷链温度监控方法[J/OL]. 农业机械学报, 2017, 48(8): 302–308. LI Zhigang, LIU Dandan, ZHANG Xiaoshuan. Cold chain temperature monitoring method of agricultural products based on clustered data fusion[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(8): 302–308. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20170835&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2017.08.035. (in Chinese)
- [2] 崔琳. 北方日光温室无线传感器多数据融合技术的研究[D]. 沈阳: 沈阳农业大学, 2016. CUI Lin. Study on data analysis platform of northern sunlight greenhouse's monitor and control system[D]. Shenyang: Shenyang Agricultural University, 2016. (in Chinese)
- [3] 熊迎军, 沈明霞, 陆明洲, 等. 温室无线传感器网络系统实时数据融合算法[J]. 农业工程学报, 2012, 28(23): 160–166. XIONG Yingjun, SHEN Mingxia, LU Mingzhou, et al. Algorithm of real time data fusion for greenhouse WSN system[J]. Transactions of the CSAE, 2012, 28(23): 160–166. (in Chinese)
- [4] SALAZAR R, LÓPEZ I, ROJANO A. A neural network model to control greenhouse environment[C] // Sixth Mexican International Conference on Artificial Intelligence-Special Session, MICAI 2007. IEEE, 2007: 311–318.
- [5] JARAMILLO V H, OTTEWILL J R, DUDEK R, et al. Condition monitoring of distributed systems using two-stage Bayesian inference data fusion[J]. Mechanical Systems and Signal Processing, 2017, 87: 91–110.
- [6] 王静. 无线传感器网络温室监测系统的设计及数据融合算法的研究[D]. 呼和浩特: 内蒙古农业大学, 2014. WANG Jing. Greenhouse monitoring system design and data fusion algorithm research based on wireless sensor network[D]. Huhhot: Inner Mongolia Agricultural University, 2014. (in Chinese)
- [7] 张品, 董为浩, 高大冬. 一种优化的贝叶斯估计多传感器数据融合方法[J]. 传感技术学报, 2014, 27(5): 643–648. ZHANG Pin, DONG Weihao, GAO Dadong. An optimal method of data fusion for multi-sensors based on Bayesian estimation[J]. Chinese Journal of Sensors and Actuators, 2014, 27(5): 643–648. (in Chinese)
- [8] 孙力帆, 张雅媛, 郑国强, 等. 基于 D-S 证据理论的智能温室环境控制决策融合方法[J/OL]. 农业机械学报, 2018, 49(1): 268–275. SUN Lifan, ZHANG Yayuan, ZHENG Guoqiang, et al. Approach to decision fusion for intelligent greenhouse environmental control based on D-S evidence theory[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(1): 268–275. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20180133&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2018.01.033. (in Chinese)
- [9] 李琼. 温室监控系统中多传感器数据融合技术的研究及应用[D]. 银川: 宁夏大学, 2013. LI Qiong. Research and application of multisensor data fusion technology in the greenhouse monitoring system[D]. Yinchuan: Ningxia University, 2013. (in Chinese)
- [10] 周观民, 李荣会. 基于神经网络的传感器网络数据融合技术研究[J]. 计算机仿真, 2011, 28(10): 118–120, 160. ZHOU Guanmin, LI Ronghui. Sensor network based on neural network data fusion technology[J]. Computer Simulation, 2011, 28(10): 118–120, 160. (in Chinese)
- [11] 孙凌逸, 黄先祥, 蔡伟, 等. 基于神经网络的无线传感器网络数据融合算法[J]. 传感技术学报, 2011, 24(1): 122–127. SUN Lingyi, HUANG Xianxiang, CAI Wei, et al. Data aggregation of wireless sensor networks using artificial neural networks[J]. Chinese Journal of Sensors and Actuators, 2011, 24(1): 122–127. (in Chinese)
- [12] 杨帆, 孟翔飞, 孙建红. 数据融合技术在温室环境监控系统中的应用[J]. 农机化研究, 2012, 34(4): 177–180. YANG Fan, MENG Xiangfei, SUN Jianhong. Applying of data fusion technology in greenhouse environment monitoring and control system[J]. Agricultural Mechanization Research, 2012, 34(4): 177–180. (in Chinese)
- [13] 黄小红. 传感器网络数据融合技术研究及在温室控制中的应用[D]. 成都: 电子科技大学, 2009.

Analytical Atomic Spectrometry, 2015, 30(2): 368–374.

- [15] 董春旺,赵杰文,朱宏凯,等. 基于 RSM 和 BP-AdaBoost-GA 的红茶发酵性能参数优化[J/OL]. 农业机械学报,2017,48(5):335–342.
DONG Chunwang, ZHAO Jiewen, ZHU Hongkai, et al. Parameter optimization of black tea fermentation machine based on RSM and BP-AdaBoost-GA[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(5):335–342. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20170542&flag=1&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2017.05.042. (in Chinese)
- [16] MALLEY D F, MCCLURE C, MARTIN P D, et al. Compositional analysis of cattle manure during composting using a field-portable near-infrared spectrometer[J]. Communications in Soil Science & Plant Analysis, 2005, 36(4–6): 455–475.
- [17] 张俊宁,方宪法,张小超,等. 基于激光诱导击穿光谱的土壤钾素检测[J/OL]. 农业机械学报,2014,45(10):294–299.
ZHANG Junning, FANG Xianfa, ZHANG Xiaochao, et al. Detection of soil potassium content based on laser-induced breakdown spectroscopy[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(10): 294–299. http://www.j-csam.org/ch/reader/view_abstract.aspx?file_no=20141046&flag=1&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2014.10.046. (in Chinese)
- [18] 张贵银,李慧,靳一东. 样品的激光诱导击穿光谱及谱线的自吸收现象[J]. 光谱学与光谱分析,2014,34(12):3397–3400.
ZHANG Guiyin, LI Hui, JIN Yidong. Laser induced breakdown spectra of coal sample and self-absorption of the spectral line [J]. Spectroscopy and Spectral Analysis, 2014, 34(12): 3397–3400. (in Chinese)
- [19] 胡慧琴,徐雪红,刘木华,等. 激光诱导击穿光谱结合偏最小二乘法直接检测皮蛋壳中的 Cu 含量[J]. 光谱学与光谱分析,2015,35(12):3500–3504.
HU Huiqin, XU Xuehong, LIU Muhua, et al. Determination of Cu in shell of preserved egg by LIBS coupled with PLS[J]. Spectroscopy and Spectral Analysis, 2015, 35(12): 3500–3504. (in Chinese)
- [20] 郭明才. 原子吸收光谱分析应用指南[M]. 青岛: 中国海洋大学出版社,2012.

(上接第 280 页)

- [14] OUTANOUTE M, LACHHAB A, ED-DAHAK A, et al. A neural network dynamic model for temperature and relative humidity control under greenhouse[C]//2015 Third International Workshop on RFID And Adaptive Wireless Sensor Networks (RAWSN). IEEE, 2015: 6–11.
- [15] 王小雪. 设施番茄高效生态栽培技术要点[J]. 南方农业,2017,11(31):61–63.
- [16] 王健. 番茄生长发育模型研究及其专家系统设计[D]. 北京:北京理工大学,2015.
WANG Jian. Research on growth and development model of tomato and its expert system design[D]. Beijing:Beijing Institute of Technology, 2015. (in Chinese)
- [17] LIU Yinghui, DENG Genqing. Study on data fusion of wireless monitoring system for greenhouse[C]//2015 8th International Conference on Intelligent Computation Technology and Automation (ICICTA). IEEE, 2015: 864–866.
- [18] WANG X H, XU L H, WEI R H. A new fusion structure model on greenhouse environment data and a new fusion algorithm of sunlight[C]//2014 International Conference on Wireless Communication and Sensor Network (WCSN). IEEE, 2014: 418–424.
- [19] CARRASQUILLA B A, CHACÓN R A, SOLÓRZANO Q M. Using IOT resources to enhance the accuracy of overdrain measurements in greenhouse horticulture[C]//2016 IEEE 36th Central American and Panama Convention (CONCAPAN XXXVI). IEEE, 2016: 1–5.
- [20] 赵树林,徐鹏民,吕光杰,等. 数据融合算法在农业物联网信息采集中的研究与应用[J]. 青岛农业大学学报(自然科学版),2016,33(1):57–60,67.
ZHAO Shulin, XU Pengmin, LÜ Guangjie, et al. Research and application of data fusion algorithm in agricultural IOT information collection[J]. Journal of Qingdao Agricultural University (Natural Science Edition), 2016, 33(1): 57–60,67. (in Chinese)
- [21] 李峰. 在农业物联网中基于卡尔曼滤波算法实现系统数据的融合处理[J]. 农业网络信息,2014(12):13–15.
LI Feng. The system data fusion in agricultural internet of things based on Kalman filtering algorithm[J]. Agricultural Network Information, 2014(12): 13–15. (in Chinese)
- [22] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436.
- [23] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527–1554.
- [24] GOODFELLOW I J, BULATOV Y, IBARZ J, et al. Multi-digit number recognition from street view imagery using deep convolutional neural networks[C]. Cornell University; arXiv: 1312.6082, 2013.
- [25] CHENG H T, KOC L, HARMSSEN J, et al. Wide & deep learning for recommender systems[C]//Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. ACM, 2016: 7–10.
- [26] 谢勇,杜建军,李永胜,等. 无公害番茄无土栽培生产技术规程[J]. 广东农业科学,2006(12):84–87.