doi:10.6041/j.issn.1000-1298.2020.02.023

基于迁移学习的 FDR 土壤水分传感器自动标定模型研究

李鸿儒[1]　于唯楚[1]　王振营[2]

(1. 东北大学信息科学与工程学院, 沈阳 110819; 2. 沈阳巍图农业科技有限公司, 沈阳 110021)

摘要: 针对频域反射技术(FDR)传感器人工标定数据拟合误差大的问题,引入其他地区数据作为辅助数据,建立了基于迁移学习的自动标定模型。该模型将 FDR 目标使用地点采集的数据作为源域数据,结合辅助数据与少量源域数据,使用 TrAdaBoost 算法即可得到准确的 FDR 传感器标定模型。将面向分类问题的 TrAdaBoost 算法改进为适用于本文面向回归的 TrAdaBoost 算法,将 TrAdaBoost 算法的基学习器由 AdaBoost 改为 XGBoost,改进了更新权重误差率的计算方法。首先使用 XGBoost 对辅助数据进行训练,得到初始标定模型;然后在目标地点采集少量数据,使用改进后的 TrAdaBoost 算法对初始标定模型进行校准,即可得到准确的 FDR 标定模型。将 10 个不同地区站点数据作为辅助数据,训练得到初始标定模型,将沈阳地区 6 个站点分别作为目标使用地点,取 80% 数据作为源域数据,进行模型校正,其余 20% 数据用于测试。测试结果的平均准确率为 99.1%,说明基于迁移学习的自动标定模型是有效和准确的。

关键词: FDR 土壤水分传感器; 自动标定; 迁移学习; TrAdaBoost; XGBoost

中图分类号: S24; S152.7　　**文献标识码**: A　　**文章编号**: 1000-1298(2020)02-0213-08　　**OSID**:

Automatic Calibration Model of FDR Soil Moisture Based on Transfer Learning

LI Hongru[1]　YU Weichu[1]　WANG Zhenying[2]

(1. *College of Information Science and Engineering*, *Northeastern University*, *Shenyang* 110819, *China*
2. *Shenyang Weitu Agricultural Science and Technology Co.*, *Ltd.*, *Shenyang* 110021, *China*)

Abstract: Aiming at the problem of large fitting error of manual calibration data for FDR sensors, the data from other regions were introduced as auxiliary data, and an automatic calibration model based on migration learning was established. In this model, historical data from other regions were introduced as auxiliary data. Data collected from FDR targets were used as source data. Combined with auxiliary data and a small amount of source data, an accurate FDR sensor calibration model can be obtained by using TrAdaBoost algorithm. TrAdaBoost algorithm for classification problem was improved to TrAdaBoost algorithm for regression. The basic learner of TrAdaBoost algorithm was changed from AdaBoost to XGBoost, which improved the calculation method of error rate when updating weight. Firstly, XGBoost was used to train the auxiliary data to get the initial calibration model, and then a small amount of data was collected from the target location of FDR, and the improved TrAdaBoost algorithm was used to calibrate the initial calibration model, so that the accurate FDR calibration model can be obtained. The data of 10 different regional sites were trained as auxiliary data to obtain the initial calibration model. For the six sites in Shenyang, the target sites were used respectively. Totally 80% of the data were used as the source domain data for model correction, and the remaining 20% were used for testing. The results showed that the average preparation rate using the calibration method was 99.1%, which indicated that the automatic calibration model using migration learning was effective and accurate.

Key words: FDR soil moisture sensor; automatic calibration; transfer learning; TrAdaBoost; XGBoost

收稿日期: 2019-05-10　修回日期: 2019-08-26

基金项目: 国家重点研发计划项目(2017YFB0304205)和国家自然科学基金项目(61533007)

作者简介: 李鸿儒(1968—),男,教授,博士生导师,主要从事人工智能和复杂系统建模、控制与优化研究,E-mail: lihongru@mail.neu.edu.cn

0 引言

土壤含水率是土壤的重要参数,也是农业灌溉决策、管理中的基础数据[1-2],准确获得可靠的土壤含水率在农业生产中极为重要。近年来,随着传感器技术的发展,利用频域反射技术(FDR)对土壤含水率进行测量,获得了业内的普遍认可,得到了广泛应用[3-5]。

土壤物理性质多种多样,不同地区以及同一地区的不同时间土壤性质也会存在差异,使得 FDR 传感器在不同土壤中的测量结果不同,因此 FDR 传感器在某区域首次使用以及使用一段时间后均需重新标定。对于传感器的标定方法已有许多学者进行了研究。文献[6]提出了土壤水分传感器的三级标定方法,即对土壤水分传感器的安装标定、田间标定以及测量数据标定,提高了测量数据的准确性,但标定过程较为繁琐。文献[7]证明了 FDR 系统适用于检测湿地土壤中的土壤含水率,但校准程序受到土壤性质的限制。文献[8]分析得出,土壤水分站点数据可用性低,是因为田间标定法在自然条件下几乎无法得到覆盖土壤各个湿度区间的均匀样本数据,导致二次标定参数不合理。综合以上研究可发现,FDR 传感器标定存在数据采集耗时、费力,可用于标定的有效数据较少,人工操作与参数拟合存在一定误差等问题。

随着科技的发展,机器学习技术在土壤含水率建模方面得到了广泛应用。文献[9]在室内利用 ASD FieldSpec 3 型高光谱仪获取土壤的原始光谱,在进行数据预处理和不同数学变换后,通过最小二乘回归法、逐步回归法、岭回归法建立了土壤含水率高光谱模型;文献[10]通过对数据进行归一化处理和数据融合,能够根据不同区域进行划分和在不同作物灌水下限进行相应的运算,从而得到估计精度较高、区域大小可调的多尺度精准灌溉决策信息;文献[11]提出电容式土壤湿度传感器大规模校准的半自动化框架,但是不能完全消除数据不确定性带来的影响。如果将不同地区数据进行统计,使用机器学习方法进行标定,则可解决标定的有效数据较少的问题,但不同地区数据特征存在差异。相关学者使用了迁移学习的解决方法,如文献[12-13]采用迁移学习方法解决了小样本下图像识别准确率低的问题;文献[14]采用深度迁移学习对柑橘叶片钾含量进行了精准预测。

针对当前 FDR 传感器标定问题,本文以沈阳地区采集到的壤土为研究对象,考虑土壤性质、温度等因素对 FDR 传感器测量结果的影响,利用其他地区已获取的数据,采用机器学习方法训练模型,实现不同目标域之间的样本迁移和融合,建立基于迁移学习的 FDR 传感器自动标定模型。

1 数据来源

1.1 研究区概况

研究区位于辽宁省沈阳市(123.4°E,41.78°N,海拔 5~441 m),该地区为温带半湿润大陆性气候,年平均气温 6.2~9.7℃。

沈阳巍图农业科技有限公司对 FDR 传感器有长期的研究基础,在全国范围建有测试站点。为本文的研究提供了沈阳地区 6 个测试站点的数据,分别为站点 11~16,其他地区 10 个站点数据分别为站点 1~10(分别为北京、天津、西安、武汉、广州、重庆、大连、哈尔滨、银川、长春)。

本文将沈阳地区各站点作为 FDR 传感器目标使用地点,其他地区数据作为参考。沈阳地区供试土壤类型为壤土,土粒密度为 2.70 g/cm^3,容重为 1.2 g/cm^3,孔隙度 55%,颗粒组成见表 1。

表 1 实验土壤颗粒组成

Tab.1 Composition of soil particles

粒径/mm	质量分数/%
0~0.002	10.3
0.002~0.075	67.5
0.075~0.250	17.3
0.250~0.850	0.2
0.850~2.000	4.7

1.2 干燥法土壤水分测量原理

干燥法也叫称量法。利用恒温箱,在温度为 105℃的条件下将土壤干燥至恒定质量,干燥前后土壤质量做差,再与干燥达恒定质量时的干土质量做比值,结果与土壤容重相乘即得到土壤体积含水率 θ,采用百分数的形式表示。

1.3 FDR 土壤水分测量原理

FDR 型土壤水分测定传感器是一种利用 LC 电路的电磁振荡,根据电磁波在不同介质中振荡频率的变化来测定介质的介电常数 ε,通过一定的对应关系反演出真实土壤体积含水率 θ_v 的仪器。传感器采用串联 LC 谐振电路,其等效电路如图 1 所示[7]。

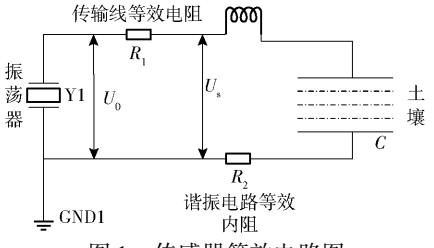

图 1 传感器等效电路图

Fig.1 Sensor equivalent circuit

根据电路原理，当谐振发生的条件成立时，谐振频率

$$F = \frac{1}{2\pi\sqrt{LC}} \tag{1}$$

式中　L——电感，H　　C——电容，F

采用新型水盐一体传感器，传感器标定过程中，首先在室内通过专用设备测试各层传感器在空气和纯水中的频率，以确定传感器的基点和极大值，用于对传感器测试结果归一化，归一化频率定义为[8]

$$F_n = \frac{F_a - F_s}{F_a - F_w} \tag{2}$$

式中　F_n——归一化频率

F_a——空气中传感器输出频率，Hz

F_w——纯水中传感器输出频率，Hz

F_s——土壤中传感器输出频率，Hz

研究发现，土壤介电常数与温度有关，使得 FDR 测得频率存在误差。针对该问题，文献[15－18]建立了温度对土壤体积含水率的补偿模型，减小了温度对传感器测量精度的影响；文献[19]通过实验的方式选择了 75 MHz 为最佳频率，消除温度对频率的影响。本文考虑在传感器设计时已加入了温度补偿模型，对此不做深入探究。

传统人工标定方法的 FDR 传感器标定经验公式为

$$\theta_v = aF_n^b \tag{3}$$

式中　a、b——标定参数

使用多组真实体积含水率 θ_v 与归一化频率 F_n 即可拟合曲线得到参数 a、b，从而得到 FDR 测量的频率与土壤体积含水率的函数关系式。

1.4　数据采集

采用沈阳巍图农业科技有限公司研制的新型水盐一体传感器对全国 16 个站点进行长时间数据采集，涵盖了土壤水分稳定期、缓慢消耗期、大量损耗期及恢复期的不同含水率的土样。在每个站点分多个土层（土层深度为测试点距地表距离分别为 10、20、40、60、80、100 cm）进行测试，测试记录数据包括站点、土层深度、传感器输出频率（包括 F_a、F_w、F_s）以及与输出频率对应的真实体积含水率（干燥法测得），测试土壤体积含水率为 5% ~50%，其中站点 1 数据如表 2 所示。

表 2　站点 1 不同土层数据

Tab. 2　Data of different soil layers in experimental site 1

土层深度/cm	F_a/Hz	F_w/Hz	F_s/Hz	F_n	体积含水率/%
10	87.089	64.475	73.724	0.591	10.627
	87.089	64.475	71.847	0.674	18.933
	86.881	61.523	72.652	0.561	19.322
	86.881	61.523	72.758	0.557	24.287
	86.881	61.523	75.110	0.464	26.200
	86.881	61.523	68.306	0.733	30.360
20	86.430	65.836	73.106	0.647	13.065
	86.994	61.290	63.354	0.920	19.750
	86.994	61.290	75.873	0.433	22.137
	86.994	61.290	78.709	0.322	25.666
	86.994	61.290	75.825	0.435	30.465
	86.994	61.290	63.354	0.920	34.855
40	86.195	65.171	73.707	0.594	6.225
	86.195	65.171	74.611	0.551	9.370
	85.671	59.684	73.072	0.485	29.180
	85.671	59.684	75.815	0.379	32.605
	85.671	59.684	79.491	0.238	34.200
	85.671	59.684	75.911	0.376	36.105
60	86.157	65.960	79.714	0.319	3.217
	86.157	65.960	79.472	0.331	5.596
	86.312	60.211	72.844	0.516	20.030
	86.312	60.211	81.962	0.167	30.701
	86.312	60.211	72.844	0.516	31.184
	86.312	60.211	75.306	0.422	37.062
80	90.002	63.351	75.490	0.545	21.070
	90.002	63.351	75.330	0.551	31.706
	90.002	63.351	74.888	0.567	35.780
	90.002	63.351	79.395	0.398	35.910
	90.002	63.351	80.808	0.345	36.011
	90.002	63.351	79.395	0.398	37.790
100	86.399	66.882	77.343	0.464	16.881
	86.399	66.882	77.128	0.475	19.107
	88.022	61.384	71.696	0.613	30.441
	88.022	61.384	71.696	0.613	31.890
	88.022	61.384	71.722	0.612	33.570
	88.022	61.384	71.195	0.632	36.874

2　研究方法

新型 FDR 土壤水盐一体传感器输出的原始信号是频率，传感器的标定即是在频率信号、土层深度和土壤含水率之间建立函数联系。同时考虑到各个测量值与土壤体积含水率之间的相关性，计算其相关性矩阵如表 3 所示，矩阵中数据表示两参数间相关性，相关系数取值范围为[0,1]，0 表示无相关，1 表示强相关。

由表 3 可知，体积含水率与其他参数均存在一定相关性，因此在建立模型时采用 F_a、F_s、F_w、FDR 归一化频率 F_n、土层深度作为输入，土壤体积含水率 θ_v 作为输出。

FDR 土壤水分传感器标定时需使用干燥法测体积含水率，代价很高，所以可用于传感器标定的有效数据较少。因此，本文考虑结合其他地区测量的大量相关数据，使用机器学习方法建模分析。传统的机器学习方法训练和测试数据同分布，通过数据

表 3 输入输出之间的相关系数

Tab. 3 Correlation between input and output

	土层深度	F_a	F_w	F_s	F_n	体积含水率
土层深度	1					
F_a	0.028 34	1				
F_w	−0.258 52	0.157 92	1			
F_s	−0.353 38	0.106 03	0.614 47	1		
F_n	0.310 79	0.074 31	−0.301 47	−0.926 40	1	
体积含水率	0.648 21	0.448 24	−0.521 75	−0.636 93	0.562 80	1

分析知,不同地区的数据分布不完全相同,故不能直接使用其他地区数据用于当前地区的传感器标定模型训练。为此,本文引入迁移学习的方法,采用 TrAdaBoost 算法在当前地区少量有效标定数据的情况下结合其他地区数据作为辅助进行模型建立。

2.1 数据预处理与分析

2.1.1 数据预处理

在数据建模前首先需进行数据清洗,去除异常数据。箱型图不受异常值的影响,能够准确稳定地描绘出数据的离散分布情况,利于数据的清洗。本文采用箱型图的方法,处理结果如图 2、表 4 所示。针对异常值识别结果,剔除不符合框图要求的数据,剔除输入数据 F_a 在 85.07 ~ 88.64 Hz 外、F_w 在 58.55 ~ 64.85 Hz 外、F_s 在 63.73 ~ 76.58 Hz 外的数据。

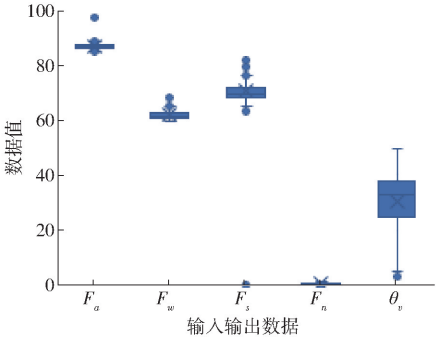

图 2 输入输出数据的箱形图

Fig. 2 Box plot of input and output data

表 4 输入输出数据排序

Tab. 4 Sort of input and output data

参数	F_a/Hz	F_w/Hz	F_s/Hz	F_n	体积含水率/%
最小值	84.84	59.37	63.35	0.167	2.80
下四分位数	86.41	60.91	68.55	0.630	24.83
中位数	86.75	61.44	69.71	0.680	32.55
上四分位数	87.30	62.49	71.76	0.720	38.01
最大值	97.37	68.87	82.88	0.919	49.87
最大值区间	88.64	64.85	76.58	0.855	57.77
最小值区间	85.07	58.55	63.73	0.495	5.07

2.1.2 数据分布分析

为验证训练数据和测试数据是否满足同分布,以站点 1 ~ 10 数据为训练数据,站点 11 ~ 16 数据作为测试数据进行分析,结果如图 3、4 所示。由图 3、4 可知,训练数据和测试数据不满足数据同分布,不能使用传统机器学习方法直接训练,故采用迁移学习方法。

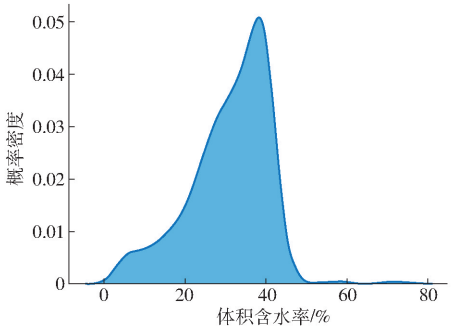

图 3 站点 1 ~ 10 训练数据概率密度分布

Fig. 3 Probability density distributions of training data at site 1 ~ 10 in other regions

2.2 基于迁移学习的自动标定模型

采用集成学习算法——基于实例的 TrAdaBoost 迁移学习算法。TrAdaBoost[20-21] 是戴文渊提出的一种基于实例的迁移学习算法,是一种从历史数据中提取实例的方法,即将一部分能用的带标签历史数据,结合带标签新数据(可能是少量),构建出比单独使用带标签新数据训练更精确的模型,适用于分类领域。本文模型的输出为体积含水率,是连续型变量,应采用回归模型,为此对 TrAdaBoost 算法进行改进,将该算法原为面向分类问题的基学习器 AdaBoost,改为面向回归问题的 XGBoost。该算法所涉及的数据集包括辅助训练数据、源域数据、目标域数据,其中辅助训练数据是指大量相关数据;源域数据是指少量与测试数据同分布的数据;目标域数据是指测试数据,即实际应用时的无标签数据。

基于其他地区 10 个站点数据(站点 1 ~ 10),沈阳地区 6 个站点数据(站点 11 ~ 16)为依据进行数据分析,将站点 1 ~ 10 的数据作为辅助训练数据(X_b,共 589 条数据,其中站点 1 为棕壤土、站点 2 为潮土、站点 3 为红粘土、站点 4 为黄棕壤土、站点 5 为红壤土、站点 6 为黄壤土、站点 7 为潮土、站点 8

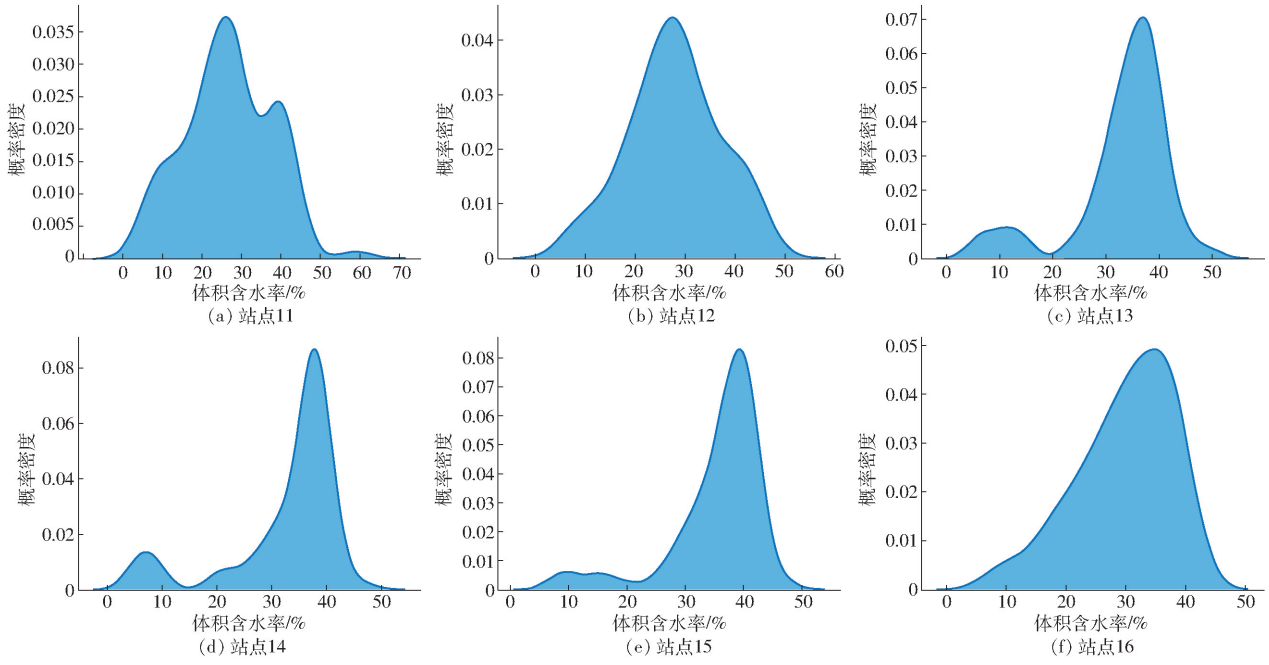

图 4　沈阳地区站点 11 ~ 16 源域概率密度分布

Fig. 4　Probability density distributions of site 11 ~ 16 source area in Shenyang

为黑土、站点 9 为灰钙土、站点 10 为黑钙土)，站点 11 ~ 16 分别作为 FDR 传感器目标使用地点，取其 80% 数据为源域数据(X_a)进行迁移学习训练得到标定模型，剩余 20% 数据作为目标域数据验证集测试模型误差。在实际应用时，源域数据的采集要求在不破坏土质的情况下涉及 10、20、40、60、80、100 cm 土层，体积含水率在 0% ~ 20%、20% ~ 30%、30% ~ 50% 均有数据，样本量最多为辅助训练数据样本的 10%，样本量最少为 20 个，否则迁移学习算法将退化为基学习器效果。

2.2.1　基学习器算法——XGBoost 模型

将辅助训练数据进行数据处理后，以 F_n、土层深度作为输入，体积含水率作为输出，利用 XGBoost 作为基学习器对其进行训练。

XGBoost[22] 是一种基于集成学习的用于处理稀疏数据的树学习算法。它的优点在于使用最少的集群资源扩展到更大的数据的端到端系统。

该算法的伪代码为：

输入：I，当前节点的实例集；d，特征维度；gain←0；$(G_L, G_R) \leftarrow \sum_{i \in I} g_i$；$(H_L, H_R) \leftarrow \sum_{i \in I} h_i$

For $k = 1$ to m：

$G_L \leftarrow 0$，$H_L \leftarrow 0$

For j in sort (I, by x_{jk}) do

$G_L \leftarrow G_L + g_j$，$H_L \leftarrow H_L + h_j$

$G_R \leftarrow G - G_j$，$H_R \leftarrow H - H_L$

$$\text{score} \leftarrow \max\left(\text{score}, \frac{G_L^2}{H_L + \lambda} + \frac{G_K^2}{H_K + \lambda} - \frac{G}{H + \lambda}\right)$$

end

end

输出：分裂后的得分 score

2.2.2　目标域——面向回归的 TrAdaBoost 算法

TrAdaBoost 算法是采用 AdaBoost 作为基学习器的分类算法，为适应本文的回归模型，将基学习器改为 XGBoost。在 TrAdaBoost 算法中对权重进行迭代更新时采用误分类样本误差率

$$c_t = \sum_i^m w_i^t I \tag{4}$$

式中　m——样本数量　　w_i^t——权重

对于回归模型，修改误差率计算式为

$$c_t = \sum_{i=n+1}^{n+m} \frac{w_i^t |h_t(x_i) - c(x_i)|}{\max |h_t(x_i) - c(x_i)|} \tag{5}$$

式中　$h_t(x_i)$——回归器的预测值

$c(x_i)$——真实值

$\max |h_t(x_i) - c(x_i)|$——训练集上样本的最大误差

改进后的 TrAdaBoost 算法既保留了 XGBoost 可降低过拟合、自动学习缺失样本的分裂方向等优点，又弥补了 XGBoost 不能进行知识迁移的缺陷。改进后的算法伪代码如下：

输入：源域数据 X_a，辅助训练数据 X_b，合并的训练数据集 $T = \{X_a \cup X_b\}$，基学习器(Learner) XGBoost，迭代次数 N。

(1)初始化

①初始化权重向量

$$\boldsymbol{w}^1 = (w_1^1, w_2^1, \cdots, w_{n+m}^1)$$

其中 $w_i^1 = \begin{cases} 1/n & (i=1,2,\cdots,n) \\ 1/m & (i=n+1,n+2,\cdots,n+m) \end{cases}$

②设置 $\beta = \dfrac{1}{1+\sqrt{2\ln(n/N)}}$

(2)权重迭代更新

对于 $t=1,2,\cdots,N$:

①设置 p^t 满足 $p^t = \dfrac{w^t}{\sum_{i=1}^{n+m} w_i^t}$

②调用 Learner,根据合并后的训练数据 T 以及 T 上的权重分布 p^t,得到一个回归器 h_t

③根据式(5)计算 h_t 在 X_b 上的误差率

④设置 $\beta_t = \dfrac{c_t}{1-c_t}$

⑤重新调整权重

$$w_i^{t+1} = \begin{cases} w_i^t \beta_t^{|h_t(x_i)-c(x_i)|} & (i=1,2,\cdots,n) \\ w_i^t \beta_t^{-|h_t(x_i)-c(x_i)|} & (i=n+1,n+2,\cdots,n+m) \end{cases}$$

⑥得到最终的回归器

(3)输出最终的回归器

对于辅助训练样本来讲,预测值和标签越接近,权重越大;而对于源域数据则相反,预测值和标签差异越大,权重越大。需要找到辅助样本中和源域数据分布最接近的样本,同时放大源域样本 loss 影响(增加源训练数据中错误率大的样本的权重,同时减小辅助训练数据中错误的权重),那么源域样本预测值与标签尽量匹配,辅助样本在前面处理的基础上筛选出最匹配(权重大的)的部分。

3 实验结果与分析

将站点 1 ~ 10 数据作为辅助训练数据,站点 11 ~ 16 分别作为源域数据,进行传感器标定模型训练与测试。对每个站点进行标定时,取该站点 80% 数据作为源域数据,共 60 条数据;剩余 20% 数据作为测试集,共 15 条数据。

根据传感器标定结果,由传感器输出频率计算得土壤体积含水率为测试值,认为干燥法得到的土壤体积含水率为真实值,使用平均百分比误差(MAPE)评估模型输出结果的准确率。

为防止过拟合,在计算测试值时,在源域数据和测试集上采用 k 折交叉验证。不重复地随机将源域数据划分为 k 个部分,$k-1$ 个部分用于训练,剩余部分用于测试,重复该过程 k 次,得到 k 个模型对模型性能的评价,基于评价结果可以计算平均性能,此方法对数据划分方法敏感度相对较低,每次迭代过程中每个样本点只有一次被划入训练集或测试集的机会。本文使用 $k=5$,即每个部分为站点数据的 20%。

为验证使用迁移学习进行初步模型校准的必要性,在站点 11 ~ 16 分别使用基学习器训练得到初步标定模型,并用迁移学习得到最终的标定模型,对比初步标定模型与最终的标定测试准确率,结果如表 5 所示,基学习器模型列为初步标定模型的准确率,自动标定模型列为使用迁移学习对初步标定模型校准后的准确率。表 5 结果显示,仅使用基学习器模型,准确率仅为 65% 左右,不能满足传感器标定要求;而使用迁移学习算法对模型进行校准后的准确率提升到了 99% 左右,已可满足传感器标定要求。故使用迁移学习的自动标定模型是有效且必要的。

表 5 标定测试准确率

Tab. 5 Calibration test accuracy %

站点编号	基学习器模型	人工标定方法	自动标定模型
11	64.30	89.75	98.29
12	66.80	88.32	99.30
13	60.23	90.83	99.34
14	61.50	92.64	99.10
15	62.48	95.86	99.15
16	63.48	94.01	99.26

针对每个站点每个土层 80% 数据使用人工标定方法计算 FDR 传感器标定参数,使用剩余 20% 数据计算人工标定传感器测量准确率,每个站点准确率为该站点测试准确率平均值,结果如表 5 所示。与本文方法进行对比发现人工标定方法准确率为 90% 左右,而本文方法平均准确率达到 99.1%,充分说明本文算法的有效性。

4 结论

(1)针对 FDR 传感器有效标定数据量少的问题,提出了基于迁移学习的 FDR 土壤水分传感器自动标定模型,仅需少量数据对模型进行校准即可使得模型输出结果满足要求。该算法克服了传统人工标定数据采集费时、费力的问题,减少了对标定数据的需求。

(2)改进了 TrAdaBoost 算法,更新了训练模型回归误差率的计算方式,并采用 k 折交叉验证有效防止了过拟合问题。

(3)将本文方法与基学习器方法的实验效果进行对比,结果显示,在使用迁移学习进行模型校准后,大大提高了模型测试准确率,说明了基于迁移学习的自动标定模型的有效性。

(4)使用本文方法,传感器在土壤体积含水率测试时,得到平均绝对误差均小于 2%,符合农业测量土壤含水率小于 ±5% 的规范要求。

参 考 文 献

[1] MILLER J D, GASKIN G J, ANDERSON H A. From drought to flood: catchment responses revealed using novel soil water probes[J]. Hydrological Processes, 2015, 11(5):533 - 541.

[2] MAO W, YANG J, ZHU Y, et al. An efficient soil water balance model based on hybrid numerical and statistical methods[J]. Journal of Hydrology, 2018, 559: 721 - 735.

[3] 冯磊,杨卫中,石庆兰,等. 基于时域传输原理的土壤水分测试仪研究[J/OL]. 农业机械学报,2017,48(3):181 - 187.
FENG Lei, YANG Weizhong, SHI Qinglan, et al. Soil moisture meter based on time domain transmission principle[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(3):181 - 187. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no=20170323&flag=1&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2017.03.023. (in Chinese)

[4] XU J, LOGSDON S D, MA X, et al. Measurement of soil water content with dielectric dispersion frequency[J]. Soil Science Society of America Journal, 2014, 78(5): 1500 - 1506.

[5] UMENYIORA C A, DRUCE R L, CURRY R D, et al. Dielectric constant of sand using TDR and FDR measurements and prediction models[J]. IEEE Transactions on Plasma Science, 2012, 40(10):2408 - 2415.

[6] 黄飞龙, 何艳丽, 陈武框. FDR 土壤水分自动站三级标定的方法[J]. 广东气象, 2011, 33(6):60 - 63.

[7] BÖHME B, BECKER M, DIEKKRÜGER B. Calibrating a FDR sensor for soil moisture monitoring in a wetland in Central Kenya[J]. Physics and Chemistry of the Earth, Parts A/B/C, 2013, 66: 101 - 111.

[8] 周旭辉. FDR 自动土壤水分数据标定问题及解决方法[J]. 气象科技, 2016, 44(4):535 - 541.

[9] 张智韬,王海峰, KARNIELI Arnon, 等. 基于岭回归的土壤含水率高光谱反演研究[J/OL]. 农业机械学报,2018,49(5):240 - 248.
ZHANG Zhitao, WANG Haifeng, KARNIELI Arnon, et al. Inversion of soil moisture content from hyperspectra based on ridge regression[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(5):240 - 248. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no=20180528&flag=1. DOI:10.6041/j.issn.1000-1298.2018.05.028. (in Chinese)

[10] 杨春曦, 刘华, 谢可心, 等. 便携式土壤湿度检测装置用于精准灌溉决策系统[J]. 农业工程学报, 2018, 34(22):84 - 91.
YANG Chunxi, LIU Hua, XIE Kexin, et al. Design of protable detection device for soil moisture and its application to precision irragation decision system[J]. Transactions of the CSAE, 2018, 34(22):84 - 91. (in Chinese)

[11] ROBERTI J A, AYRES E, LOESCHER H W, et al. A robust calibration method for continental-scale soil water content measurements[J/OL]. Vadose Zone J, 2018, 17(1):170177. DOI:10.2136/vzj2017.10.0177.

[12] 龙满生, 欧阳春娟, 刘欢, 等. 基于卷积神经网络与迁移学习的油茶病害图像识别[J]. 农业工程学报, 2018, 34(18):194 - 201.
LONG Mansheng, OUYANG Chunjuan, LIU Huan, et al. Image recognition of *Camellia oleifera* diseases based on convolutional neural network & transfer learning[J]. Transactions of the CSAE, 2018, 34(18):194 - 201. (in Chinese)

[13] 郑一力,张露. 基于迁移学习的卷积神经网络植物叶片图像识别方法[J/OL]. 农业机械学报,2018,49(增刊):354 - 359.
ZHENG Yili, ZHANG Lu. Plant leaf image recognition method based on transfer learning with convolutional neural networks [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(Supp.):354 - 359. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=2018s047&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2018.S0.047. (in Chinese)

[14] 岳学军, 凌康杰, 王林惠, 等. 基于高光谱和深度迁移学习的柑橘叶片钾含量反演[J/OL]. 农业机械学报, 2019, 50(3):186 - 195.
YUE Xuejun, LING Kangjie, WANG Linhui, et al. Inversion of potassium content for citrus leaves based on hyperspectral and deep transfer learning[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(3): 186 - 195. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20190320&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.03.020. (in Chinese)

[15] HILHORST M A, VAN BREUGEL K, PLUIMGRAAFF D J M H, et al. Dielectric sensors used in environmental and construction engineering[J]. MRS Proceedings, 1995, 411:401.

[16] 张荣标,刘骏,张磊,等. EC-5 土壤水分传感器温度影响机理及补偿方法研究[J]. 农业机械学报,2010,41(9):168 - 172.
ZHANG Rongbiao, LIU Jun, ZHANG Lei, et al. Temperature effect mechanism and compensation method of EC-5 soil moisture sensor[J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(9): 168 - 172. (in Chinese)

[17] BOGENA H R, HUISMAN J A, OBERDÖRSTER C, et al. Evaluation of a low-cost soil water content sensor for wireless network applications[J]. Journal of Hydrology, 2007, 344(1):32 - 42.

[18] YE Zhijie, HONG Tiansheng, JOSEPH Mwape Chileshe, et al. Multi-factor evaluation and modeling correction of EC-5 and 5TE soil moisture content sensors[J]. Transactions of the CSAE, 2012, 28(1): 157 - 166.

[19] 许景辉, 马孝义, SALLY D Logsdon, 等. 基于土壤 LF-UHF 频段介电特性的 FDR 测量频率研究[J/OL]. 农业机械学

报, 2013, 44(7):67 - 72.
XU Jinghui, MA Xiaoyi, SALLY D Logsdon, et al. Frequency measurement of FDR based on soil dielectric spectrum in LF - UHF[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(7):67 - 72. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20130713&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2013.07.013. (in Chinese)
[20] YAO Y, DORETTO G. Boosting for transfer learning with multiple sources[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010: 1855 - 1862.
[21] 戴文渊. 基于实例和特征的迁移学习算法研究[D]. 上海: 上海交通大学, 2009.
DAI Wenyuan. Instance-based and feature-based transfer learning[D]. Shanghai: Shanghai Jiao Tong University, 2009. (in Chinese)
[22] CHEN T, GUESTRIN C. Xgboost: a scalable tree boosting system[C]//Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining. ACM, 2016: 785 - 794.

(上接第 202 页)

[29] DUCHEMIN B, MAISONGRANDE P, BOULET G, et al. A simple algorithm for yield estimates: evaluation for semi-arid irrigated winter wheat monitored with green leaf area index[J]. Environmental Modelling and Software, 2008, 23(7):876 - 892.
[30] CLAVERIE M, VALERIE D, BENOIT D, et al. Maize and sunflower biomass estimation in southwest France using high spatial and temporal resolution remote sensing data[J]. Remote Sensing of Environment, 2012, 124(6):844 - 857.
[31] CLAVERIE M, VALERIE D, BENOIT D, et al. Spatialization of crop leaf area index and biomass by combining a simple crop model SAFY and high spatial and temporal resolutions remote sensing data[C]//Geoscience & Remote Sensing Symposium. IEEE, 2010.
[32] CHU W, GAO X, SOROOSHIAN S. A new evolutionary search strategy for global optimization of high-dimensional problems [J]. Information Sciences, 2011, 181(22):4909 - 4927.
[33] CHU W, GAO X, SOROOSHIAN S. A solution to the crucial problem of population degeneration in high-dimensional evolutionary optimization[J]. IEEE Systems Journal, 2011, 5(3):362 - 373.
[34] DUAN Q Y, GUPTA V K, SOROOSHIAN S. Shuffled complex evolution approach for effective and efficient global minimization[J]. Journal of Optimization Theory and Applications, 1993, 76(3):501 - 521.
[35] 郑阳, 吴炳方, 张淼. Sentinel - 2 数据的冬小麦地上干生物量估算及评价[J]. 遥感学报, 2017, 21(2):318 - 328.
ZHENG Yang, WU Bingfang, ZHANG Miao. Estimating the above ground biomass of winter wheat using the Sentinel - 2 data [J]. Journal of Remote Sensing, 2017, 21(2):318 - 328. (in Chinese)
[36] 李卫国, 顾晓鹤, 王尔美, 等. 基于作物生长模型参数调整动态估测夏玉米生物量[J]. 农业工程学报, 2019, 35(7): 136 - 142.
LI Weiguo, GU Xiaohe, WANG Ermei, et al. Dynamic estimation of summer maize biomass based on parameter adjustment of crop growth model[J]. Transactions of the CSAE, 2019, 35(7): 136 - 142. (in Chinese)
[37] 牛庆林, 冯海宽, 杨贵军, 等. 基于无人机数码影像的玉米育种材料株高和 LAI 监测[J]. 农业工程学报, 2018, 34(5): 73 - 82.
NIU Qinglin, FENG Haikuan, YANG Guijun, et al. Monitoring plant height and leaf area index of maize breeding material based on UAV digital images[J]. Transactions of the CSAE, 2018, 34(5): 73 - 82. (in Chinese)
[38] 张春兰, 杨贵军, 李贺丽, 等. 基于随机森林算法的冬小麦叶面积指数遥感反演研究[J]. 中国农业科学, 2018, 51(5):855 - 867.
ZHANG Chunlan, YANG Guijun, LI Heli, et al. Remote sensing inversion of winter wheat leaf area index based on random forest algorithm[J]. Scientia Agricultura Sinica, 2018, 51(5):855 - 867. (in Chinese)