一种隐私保护的联邦学习框架*

杨东宁[1,2],谢潇睿[1],吉志坤[3],姬维维[3]
(1.云南电网有限责任公司 信息中心,云南 昆明 650011;2.西南林业大学 大数据与智能工程学院,云南 昆明 650224;
3.云南云电同方科技有限公司,云南 昆明 650220)

摘 要:大数据时代,数据安全性和隐私性受到越来越多的关注和重视。联邦学习被视为是一种隐私保护的可行技术,允许从去中心化的数据中训练深度模型。针对电力投资系统中各部门因担心数据隐私信息泄露而带来的数据孤岛和隐私保护问题,提出了一种隐私保护的联邦学习框架,允许各部门自有数据在不出本地的情况下,联合训练模型。首先,提出了联邦学习的架构,支持分布式地训练模型;其次,引入同态加密技术,提出了隐私保护的联邦平均学习流程,在数据隐私保护的情况下,实现联合训练模型;最后,实验结果表明,该框架具有较好的收敛性,而且联合训练得到的模型具有较好的精度。

关键词:数据隐私;联邦学习;深度学习;同态加密技术;卷积神经网络
中图分类号:TN711.1;TP311 **文献标识码**:A DOI:10.16157/j.issn.0258-7998.211828

中文引用格式:杨东宁,谢潇睿,吉志坤,等. 一种隐私保护的联邦学习框架[J].电子技术应用,2022,48(5):94-97,103.
英文引用格式:Yang Dongning,Xie Xiaorui,Ji Zhikun,et al. A privacy-preserving federated learning framework[J]. Application of Electronic Technique,2022,48(5):94-97,103.

A privacy-preserving federated learning framework

Yang Dongning[1,2],Xie Xiaorui[1],Ji Zhikun[3],Ji Weiwei[3]
(1.Information Center,Yunnan Power Grid Co.,Ltd.,Kunming 650011,China;
2.School of Big Data and Intelligent Engineering,Southwest Forestry University,Kunming 650224,China;
3.Yunnan Yundian Tongfang Technology Co.,Ltd.,Kunming 650220,China)

Abstract: In the era of big data, more and more attention has been paid to data security and privacy. Federated learning is regarded as a promising privacy-preserving technology, which allows training a deep model from decentralized data. To solve the problem of isolated data island and privacy protection caused by the fear of data privacy information leakage in the power investment system, this paper proposes a privacy-preserving federal learning framework, which allows departments to jointly train the model without releasing their local data. Firstly, a federated learning architecture is proposed to support distributed training model. Secondly, homomorphic encryption technology is introduced and a federal average learning process is proposed to realize the joint training model while the privacy of data are protected. Finally, the experimental results show that the framework has good convergence and the joint training model has good accuracy.

Key words: data privacy; federated learning; deep learning; homomorphic encryption; convolutional neural network

0 引言

大数据时代,数据的安全性和隐私性受到了越来越多的重视和关注[1]。欧盟及我国都分别相继通过了《一般数据法案》《欧洲数据保护通用条例》《中华人民共和国网络安全法》等相关法案,以保护数据的安全和隐私。

南方电网公司“十四五”数字化规划明确提出:电力智慧投资将以项目储备库为基础,通过输入约束条件和投资分配参数,并结合规划部门的配网规划数据、基建部门的基建结算数据和财务部门的财务数据,运用机器学习训练生成的投资预测模型,自动生成投资计划项目及费用估算建议。但是,目前的电力投资系统依赖集中式的方式训练模型,要求训练涉及的多方数据和训练过程必须在数据中心。在此过程中,各部门自有数据中的隐私信息可能会被泄露。此外,各部门出于数据安全和隐私保护的需求,不可能将自有数据上传到数据中心。因此,如何在确保各方数据安全和隐私的情况下打破数据孤岛、共同训练模型,成为了急需解决的挑战。

联邦学习使得机器学习或深度学习算法能从不同组织或部门的大量数据中获得更好的经验[2]。这种技术允许多个组织或部门在数据不直接共享的情况下协作

*基金项目:国家自然科学基金项目(61702442)

完成模型的联合训练[3]。具体来讲，各组织或部门的私有自有数据可以不离开本地，通过本地模型参数的更新和全局模型参数的聚合，在确保各自数据隐私性和安全性的情况下，联合训练一个共享的全局模型。因此，联邦学习被视为解决数据孤岛和打破数据壁垒的有效可行技术[4]。

为了解决电力投资系统面临的数据孤岛和隐私保护问题，本文提出了一个隐私保护的联邦学习框架，在各部门数据不出本地的情况下，联合训练投资预测模型。具体地，本文主要贡献如下：

(1)提出了一种基于客户端-服务器的联邦学习架构。与传统集中式的模型训练方式相比，该架构支持在隐私保护的情况下，分布式地联合训练模型。

(2)提出了隐私保护的联邦平均学习流程，引入同态加密技术，防止参数聚合过程实施成员推理攻击。该流程主要包括4个阶段：局部模型训练、参数加密、参数聚合和局部模型更新。

(3)实验结果表明所提方法具有较好的收敛性，而且联合训练得到的模型具有较好的精度。

1 相关工作

Lim等人[5]从通信代价、资源分配、安全性和隐私性四方面对联邦学习进行了综述。进一步，Yang等人[6]一方面从安全多方计算、差分隐私和同态加密三方面讨论了安全的联邦学习框架；另一方面，将安全的联邦学习框架划分为3种类型：水平联邦学习、垂直联邦学习和迁移联邦学习。王健宗等人[4]重点从通信负载、异步聚合等方面讨论了联邦学习的优化算法。

此外，其他工作研究了联邦学习在不同场景中的应用。Liu等人[3]将联邦学习应用于交通预测领域，提出了一种基于联邦学习的隐私保护的交通预测方法。Ye等人[7]将联邦学习和边缘计算相结合，提出了一种基于边缘计算的优化联邦学习方法。Yu等人[8]将联邦学习应用于边缘计算环境中的内容缓冲，提出了一种车联网环境下基于点对点联合学习的主动式内容缓存方法。进一步，Yu等人[9]在考虑车辆移动性和缓存内容的过期性的情况下，提出了一种基于联合学习的移动感知主动边缘缓存方案。Kim等人[10]将联邦学习和区块链相结合，提出了一种基于区块链的节点感知动态加权方法，用于提高联邦学习的性能。

与上述工作相比，本文工作主要关注将联邦学习应用于电力行业，解决电力投资系统面临的数据孤岛和隐私保护问题。此外，本文将同态加密技术引入联邦平均学习流程，用于保护上述模型参数，以防止实施成员推理攻击。

2 架构和问题定义

2.1 架构

目前，联邦式学习的架构分为两种：客户端-服务器架构[7]和对等网络架构[11]。本文提出了一种基于客户端和服务器的联邦学习架构，如图1所示。该架构包含：多个部门及部门自有的数据库、多个部门的局部模型、1个中心服务器和1个全局模型。

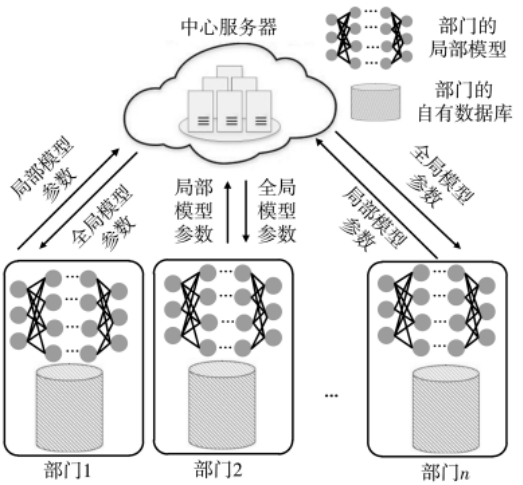

图1 面向电力投资系统的联邦学习架构

与传统集中式训练模型相比，图1所示的架构允许分布式地联合训练模型。具体而言，首先，各部门使用自有数据库在本地训练自己的局部模型；其次，将局部模型的参数经加密后上传汇总到中心服务器进行参数聚合计算，更新全局模型；最后，各部门将更新后的全局模型参数下载到本地，并更新本地模型。直至本地模型的性能收敛并足够好，分布式的训练模型过程结束。

由图1可以看出，分布式的模型训练方式通过聚合局部模型参数和使用全局模型参数更新局部模型，允许各部门的自有数据在不离开本地的情况下训练模型，既保护了各部门的数据，又共同联合训练了模型。

2.2 问题定义

对于图1所示的联邦学习架构，令$\mathcal{O}=\{O_1,O_2,\cdots,O_n\}$表示部门的集合，$\mathcal{D}=\{D_1\cup D_2\cup\cdots\cup D_n\}$表示各部门自有数据库的并集。

定义1(集中式训练模型) 给定部门集合$\mathcal{O}$和数据集$\mathcal{D}$，集中式训练得到的模型记为$M_c\leftarrow f(D)$，其中，$f(\cdot)$表示学习函数。

定义2(联邦式训练模型) 给定部门集合$\mathcal{O}$和各部门自有数据库$(D_1,D_2,\cdots,D_n)$，联邦式训练得到的模型记为$M_f\leftarrow\sum_{i\in D}w_i f_i(D_i)$，其中$w_i$表示权重，$f_i(\cdot)$表示学习函数$f(\cdot)$的本地版本。

定义3(ε-精度损失)[6] 令集中式模型M_c的预测精度为P_c，联邦式模型M_f的预测精度为P_f，若满足式(1)，则称联邦式学习算法达到ε-精度损失。

$$|P_c-P_f|<\delta \tag{1}$$

3 隐私保护的联邦平均学习流程

基于图 1 所示的架构，提出了隐私保护的联邦平均学习流程，包括 4 个步骤：局部模型训练、参数加密、参数聚合和局部模型更新。

3.1 局部模型训练

在各部门同意联合训练模型会，中心服务器将向各部门发布模型的初始参数。各部门使用自有数据对初始化后的模型进行本地训练。训练接收后，将本地训练计算得到的模型参数梯度进行上传。

对于部门 O_i，令 $\{\boldsymbol{x}_j, y_j\} \in D_i$，其中 $\boldsymbol{x}_i \in \mathbb{R}^{\mathrm{l}}$ 表示输入样本向量具有 d 个特征，其中 $y_j \in \mathbb{R}$ 表示 $\boldsymbol{x}_i$ 对应的标记输出值。对于局部模型训练，如果输入样本向量 $\boldsymbol{x}_j$，则希望从 D_i 中学习到模型参数向量 $\boldsymbol{\omega} \in \mathbb{R}^{\mathrm{l}}$，使得 $\boldsymbol{x}_j^{\mathrm{T}}\boldsymbol{\omega}$ 尽可能接近 y_j。为此，代价函数定义如下：

$$L_i(\boldsymbol{\omega}) = \frac{\sum_{\boldsymbol{x}_j \in D_i} 0.5(\boldsymbol{x}_j^{\mathrm{T}}\boldsymbol{\omega} - y_j)}{|D_i|} + \lambda h(\boldsymbol{\omega}) \tag{2}$$

其中，$\lambda \in [0,1]$，$h(\cdot)$ 表示正则化函数，$|D_i|$ 表示部门 O_i 自有数据库的条目数量。

3.2 参数加密

已有研究表明[12]：在联邦学习中，好奇的中心服务器通过共享的参数梯度，可以实施成员推理攻击，并获得训练数据信息。为了防止成员推理攻击，本文采用了同态加密技术对上传的梯度参数进行加密。

与差分隐私保护[13]和安全多方计算[14]相比，同态加密技术分别具有数据不失真和计算复杂性小的优点。由于参数聚合过程只涉及加法和乘法，将使用无需多项式近似的加法同态加密技术对梯度参数进行加密。其加密的基本原理如下：

$$D(k_{\mathrm{s}}, E(k_{\mathrm{p}}, \boldsymbol{\omega}_1) \times E(k_{\mathrm{p}}, \boldsymbol{\omega}_2)) = \boldsymbol{\omega}_1 \circ \boldsymbol{\omega}_2 \tag{3}$$

其中，$\boldsymbol{\omega}_1$、$\boldsymbol{\omega}_2$ 分别表示部门 1 和部门 2 上传的梯度参数，k_{s}、k_{p} 分别表示私钥与公钥，$E()$、$D()$ 分别表示加密运算和解密算，$\circ$、$\times$ 分别表示明文域和密文域的运算。

3.3 参数聚合

中心服务器将接收到来自各个部门的局部模型参数梯度，并对这些参数梯度进行聚合。具体而言，参数的聚合过程如下：

$$\arg\min_{\boldsymbol{\omega} \in R^{\mathrm{d}}} J(\boldsymbol{\omega}), J(\boldsymbol{\omega}) = \sum_{i=1}^{n} \frac{L_i(\boldsymbol{\omega})}{n} \tag{4}$$

根据式(2)，式(3)可进一步重写为：

$$\arg\min_{\boldsymbol{\omega} \in R^{\mathrm{d}}} J(\boldsymbol{\omega}), J(\boldsymbol{\omega}) = \sum_{i=1}^{n} \frac{\frac{\sum_{\boldsymbol{x}_j \in D_i} 0.5(\boldsymbol{x}_j^{\mathrm{T}}\boldsymbol{\omega} - y_j)}{|D_i|} + \lambda h(\boldsymbol{\omega})}{n} \tag{5}$$

3.4 局部模型更新

局部模型更新是指各部门从中心服务器下载聚合后的全局模型参数梯度，并将解密后的参数梯度用于本地局部模型的更新。具体地，对于部门 O_i，更新本地模型参数的计算如下：

$$\boldsymbol{\omega} \leftarrow \boldsymbol{w}^g \tag{6}$$

其中，$\boldsymbol{w}^g$ 表示解密后得到的全局模型参数梯度。

3.5 算法

目前，常见的联邦学习算法是谷歌公司提出的 FedAvg 算法(Federated Averaging)[15]。基于 FedAvg 算法，引入同态加密算法，提出了隐私保护的联邦平均学习算法，记为 PP-FedAvg，如算法 1 所示。

算法 1：联邦平均学习算法 PP-FedAvg

输入：部门 k 的自有数据库 D_k，b 是本地 Batch 的大小，E 表示 epoch 的数量，α 表示学习率，$\nabla L(\cdot,\cdot)$ 表示梯度优化函数；

输出：更新后的局部模型参数 $\boldsymbol{\omega}$。

(1)初始化全局模型参数 ω^0

(2)for 轮数 $t=1,2,\cdots$ do

(3) $\{\mathcal{O}_{\mathrm{v}}\} \leftarrow$从 $\mathcal{O}$ 中选取参与联邦学习的部门

(4) 将 ω^0 广播给 $\{\mathcal{O}_{\mathrm{v}}\}$ 中的每个部门

(5) for 部门 $k \in \{\mathcal{O}_{\mathrm{v}}\}$ in parallel do

(6) $\omega_{t+1}^{k} \leftarrow \mathrm{LocalUpdate}(k, \boldsymbol{\omega}_t)$

(7) end for

(8)end for

(9)输出 $\boldsymbol{\omega}_{t+1} \leftarrow \sum_{k=1}^{n} \frac{\beta_k \boldsymbol{\omega}_{t+1}}{n}$

其中，算法 1 中第 6 行调用的算法 $\mathrm{LocalUpdate}(k, \boldsymbol{\omega}_t)$ 如算法 2 所示。

算法 2：LocalUpdate 局部模型更新

输入：部门 k 的自有数据库 D_k，b 是本地批次的大小(Batch)，E 表示迭代周期的数量(epoch)，α 表示学习率，$\nabla L(\cdot,\cdot)$ 表示梯度优化函数；

输出：局部模型的参数 $\boldsymbol{\omega}_t$。

(1)for $e \in E$ do

(2) $\pi \leftarrow$将 D_k 按 b 大小分割

(3) for $b \in \pi$ do

(4) $\boldsymbol{\omega}_t \leftarrow \boldsymbol{\omega}_t - \alpha \nabla L(\boldsymbol{\omega}_t, b)$

(5) end for

(6)end for

(7)输出更新后的参数 $\boldsymbol{\omega}_t$

需要注意的是，在算法 1 中，当每轮参数聚合后，中心服务器需要确定是否需要继续执行联邦平均学习算法。

4 实验

4.1 实验设置

实验环境考虑了由 1 个中心服务器和 10 个部门组成的联邦学习架构。当从中心服务器下载初始模型参数后，10 个部门在本地并行训练模型，并将训练得到的模型参数同步发送到中心服务器进行全局聚合。这被视为

一轮通信的结束。

实验选择开源数据集 MNIST 作为数据集。该数据集包含 60 000 个训练图像和 10 000 个测试图像。其中，每个图像的分辨率为 28×28，对应于 10 个(0~9)可能数字中的 1 个。为了确保每个部门数据集满足非独立同分布，根据数字大小，将 60 000 个训练数据排序并划分为 20 份(每份 300 个图像)，并随机地分配 2 份给 10 个部门。

实验采用的编程语言为 Python，使用的深度学习模型为卷积神经网络(CNN)，如图 2 所示。该模型的输入是 28×28 的图像，输出是数字。具体地，模型由 2 种类型卷积层、2 个最大池化层和 1 个全连接层组成。第一种类型的卷积层包含两层卷积(层 2 和 3)，分别使用 16 个 3×3 的卷积核，步长为 1。第二种类型的卷积层包含两层卷积(层 5 和 6)，分别使用 32 个 3×3 的卷积核，步长为 1。两个最大池化层均使用 2×2 的卷积核。全连接层包含 10 个神经单元。

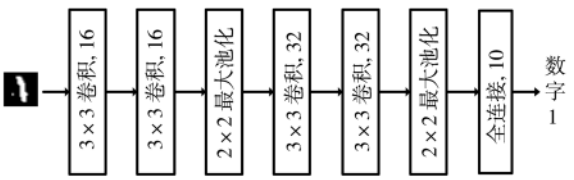

图 2 卷积神经网络的结构

为了验证所提隐私保护联邦学习框架的有效性，将框架分布式联合训练模型的方法(记为 PFedAvg)与集中式训练模型的方法(记为 CTCNN)做对比。

4.2 实验结果与分析

图 3 显示了 PFedAvg 和 CTCNN 训练模型的损失。从结果来看，用 PFedAvg 训练模型的损失与用 CTCNN 训练模型的损失没有显著差异。这不仅表明 PFedAvg 训练的模型具有良好的收敛性，还表明 PFedAvg 训练的模型具有较好的精度。

图 4 比较了在不同的批次大小和迭代周期，PFedAvg 训练模型的精度随通信轮数的变化趋势。从结果来看，

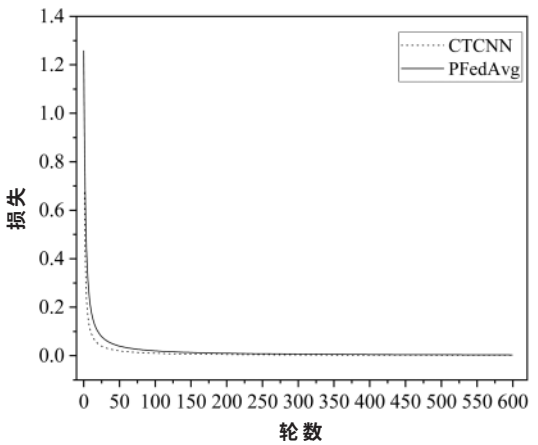

图 3 PFedAvg 和 CTCNN 训练模型的损失

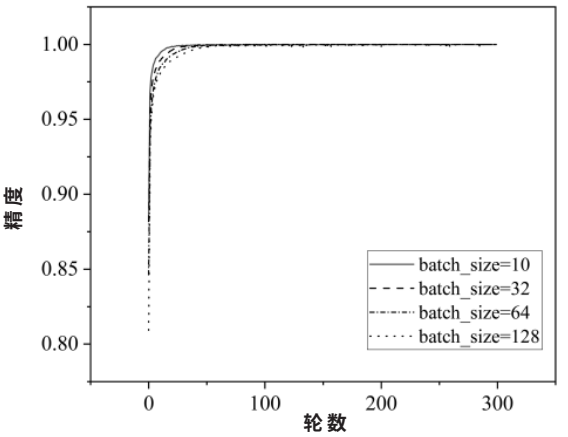

图 4 在不同批次大小和迭代周期下 PFedAvg 模型的精度

当批次大小设置为 10、迭代周期为 12 时，PFedAvg 训练得到的模型精度最高。

5 结论

为了解决电力投资系统面临的数据孤岛和隐私保护问题，本文将联邦学习应用于电力行业，提出了一种隐私保护的联邦学习框架。实验结果表明了该框架的有效性。

参考文献

[1] 叶青青，孟小峰，朱敏杰，等.本地化差分隐私研究综述[J].软件学报，2018，29(7)：1981-2005.

[2] IEEE SA.IEEE guide for architectural framework and application of federated machine learning[S].IEEE Std 3652.1-2020，2020.

[3] LIU Y，YU J J Q，KANG J，et al.Privacy-preserving traffic flow prediction: a federated learning approach[J].IEEE Internet of Things Journal，2020，7(8)：7751-7763.

[4] 王健宗，黄章成，陈霖捷，等.联邦学习算法综述[J].大数据，2020，6(6)：19.

[5] LIM W Y B，LUONG N C，HOANG D T，et al.Federated learning in mobile edge networks：a comprehensive survey[J].IEEE Communications Surveys & Tutorials，2020，22(3)：2031-2063.

[6] YANG Q，LIU Y，CHEN T，et al.Federated machine learning：concept and applications[J].ACM Trans.Intell.Syst.Technol.，2019，10(2)：12.

[7] YE Y，LI S，LIU F，et al.EdgeFed：optimized federated learning based on edge computing[J].IEEE Access，2020，8：209191-209198.

[8] YU Z，HU J，MIN G，et al.Proactive content caching for Internet-of-Vehicles based on peer-to-peer federated learning[C]//2020 IEEE 26th International Conference on Parallel and Distributed Systems(ICPADS)，2020：601-608.

[9] YU Z，HU J，MIN G，et al.Mobility-aware proactive edge caching for connected vehicles using federated learning[J].

(下转第 103 页)

征，使用空间插值法绘制脑电地形图，搭建 LeNet-5 并进行训练，在情绪的分类识别上取得了良好的分类效果，有较高的可行性。

5 结论

为了实现脑电信号情绪识别，本文采用了对脑电信号滤波、EEMD 分解提取信号特征和卷积神经网络分类方法，得到了情绪与 fp1、fp2、f3、f4、f7、f8 6 个导联 β 波信号的具体关系。实验表明，前额叶 β 波信号可以反映人的情绪，所以研究前额叶 β 波信号有利于基于脑电信号的情绪识别进行，相较于其他特征功率谱密度可以较准确地进行情绪识别。

本实验通过 EEMD 提取的前额叶 β 波脑电信号的 6 类特征值信号特征搭建 LeNet-5，选择 DEAP 数据库中的数据进行训练，在情绪的分类识别上取得了良好的分类效果，6 种特征值间相互比较，提高了信号分类准确率，识别准确率最高为功率谱密度为 80.1%，有较高的可行性。

参考文献

[1] 晁浩，刘永利，连卫芳.EEG 情感识别中基于集成深度学习模型的多分析域特征融合[J].控制与决策，2020，35(7)：1674-1680.
[2] 陆文娟.基于脑电信号的情感识别研究[D].南京：南京邮电大学，2017.
[3] 金雨鑫，骆懿，于洋.基于深度森林的脑电情绪识别研究[J].软件导刊，2019，18(7)：53-55，59.
[4] 曾红梅.情绪图片视觉诱发 EEG 特征提取与分析[D].天津：天津大学，2012.
[5] 李明爱，张梦，孙炎珺.基于小波包和深度信念网络的脑电特征提取方法[J].电子测量与仪器学报，2018，32(1)：111-118.
[6] 柳素红，孙晓，李春彬.基于位置信息重建与时频域信息融合的脑电信号情感识别[J].计算机工程，2021，47(12)：95-102.
[7] 陈萌，李幼军，刘岩.脑电信号与个人情绪状态关联性分析研究[J].计算机科学与探索，2017，11(5)：794-801.
[8] 梅寒.基于脑电导联空间关联性的情绪识别算法框架研究[D].广州：华南理工大学，2018.
[9] 张琪.基于情绪认知的视听诱发脑电信号处理与分析[D].太原：山西大学，2016.
[10] 王婷.EMD 算法研究及其在信号去噪中的应用[D].哈尔滨：哈尔滨工程大学，2010.
[11] 胡爱军，孙敬敬，向玲.经验模态分解中的模态混叠问题[J].振动·测试与诊断，2011，31(4)：429-434，532-533.
[12] 霍延.基于 EEMD 与改进 EMD 的脑电信号的特征提取方法[D].南京：南京邮电大学，2020.
[13] 谢佳利.多尺度熵算法在情感脑电识别中的应用[D].秦皇岛：燕山大学，2016：32-33.
[14] 张恒，黄莺，刘明宏，等.基于空间插值法的遵义烟区植烟土壤养分时空变化[J].中国烟草科学，2020，41(3)：36-43.
[15] 王济民，魏怡，周宇，等.基于 LeNet-5 卷积神经网络和颜色特征的限速标志识别[J].计算机科学，2021，48(S2)：345-350.

(收稿日期：2021-05-07)

作者简介：
蔡靖(1979-)，男，硕士，高级工程师，主要研究方向：医疗仪器。
孙慧慧(1984-)，通信作者，女，硕士，工程师，主要研究方向：控制系统与数据处理，E-mail：sunhuihui@jlu.edu.cn。

扫码下载电子文档

(上接第 97 页)

IEEE Transactions on Intelligent Transportation Systems，2021，22(8)：5341-5351.
[10] KIM Y J，HONG C S.Blockchain-based node-aware dynamic weighting methods for improving federated learning performance[C]//2019 20th Asia-Pacific Network Operations and Management Symposium(APNOMS)，2019：1-4.
[11] LIU X，LI H，XU G，et al.Adaptive privacy-preserving federated learning[J].Peer-to-Peer Networking and Applications，2020，13：2356-2366.
[12] NASR M，SHOKRI R，HOUMANSADR A.Comprehensive privacy analysis of deep learning：passive and active white-box inference attacks against centralized and federated learning[C]//2019 IEEE Symposium on Security and Privacy(SP)，2019：739-753.
[13] DWORK C，LEI J.Differential privacy and robust statistics[C]//Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing，2009：371-380.
[14] LINDELL Y，PINKAS B.A proof of security of Yao′s protocol for two-party computation[J].Journal of Cryptology，2009，22(2)：161-188.
[15] MILLS J，HU J，MIN G.Communication-efficient federated learning for wireless edge intelligence in IoT[J].IEEE Internet of Things Journal，2020，7(7)：5986-5994.

(收稿日期：2021-06-02)

作者简介：
杨东宁(1986-)，男，硕士，高级工程师，主要研究方向：电网信息化。

扫码下载电子文档

版权声明