计算机科学
COMPUTER SCIENCE

面向公平性联邦学习的指纹识别算法

王晨卓, 鲁艳蓉, 沈剑

引用本文

王晨卓, 鲁艳蓉, 沈剑. 面向公平性联邦学习的指纹识别算法[J]. 计算机科学, 2024, 51(6A): 230800043-9.

WANG Chenzhuo, LU Yanrong, SHEN Jian. Study on Fingerprint Recognition Algorithm for Fairness in Federated Learning [J]. Computer Science, 2024, 51(6A): 230800043-9.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.230800043

面向公平性联邦学习的指纹识别算法

王晨卓[1]　鲁艳蓉[1,2]　沈　剑[3]
1 中国民航大学计算机科学与技术学院　天津 300300
2 中国民航大学安全科学与工程学院　天津 300300
3 浙江理工大学信息科学与工程学院　杭州 310018
(2021051015@cauc.edu.cn)

摘　要　现有的指纹识别方法大多是基于机器学习,在对海量数据集中训练时忽视了数据本身的隐私性和异质性,从而导致用户信息泄漏和识别率降低。为在隐私保护下协同优化模型精度,提出了一个全新的基于联邦学习的指纹识别算法(Federated Learning-Fingerprint Recognition,Fed-FR)。首先,通过联邦学习迭代聚合来自各终端的参数,从而提高全局模型的性能;其次,将稀疏表示理论用于低质量指纹图像去噪处理,来增强指纹的纹理结构;再次,针对客户端异构而导致的分配不公问题,提出基于水库抽样的客户端调度策略;最后,在3个真实数据集上进行仿真实验,对Fed-FR的有效性进行对比分析。实验结果表明,Fed-FR精度比局部学习提高5.32%,比联邦平均算法提高8.56%,接近于集中学习的精度;在隐私保护水平、评估准确率及可扩展性等方面具有良好的表现。研究成果首次展现了联邦学习与指纹识别结合的可行性,增强了指纹识别算法的安全性和可扩展性,给联邦学习应用于生物识别技术提供了参考。

关键词:指纹识别;联邦学习;稀疏表示;水库抽样;隐私保护
中图分类号　TP391

Study on Fingerprint Recognition Algorithm for Fairness in Federated Learning

WANG Chenzhuo[1],LU Yanrong[1,2] and SHEN Jian[3]
1 School of Computer Science and Technology,Civil Aviation University of China,Tianjin 300300,China
2 School of Safety Science and Engineering,Civil Aviation University of China,Tianjin 300300,China
3 School of Informatics Science and Engineering,Zhejiang Sci-Tech University,Hangzhou 310018,China

Abstract　Most existing fingerprint recognition methods rely on machine learning,which neglects the privacy and heterogeneity of the data when training on massive databases,resulting in user information leakage and reduced recognition accuracy. To cooperatively optimize model accuracy under privacy protection,this paper proposes a novel fingerprint recognition algorithm based on federated learning,termed federated learning-fingerprint recognition(Fed-FR). Firstly,the algorithm iteratively aggregates parameters from each terminal through federated learning,thereby improving the performance of the global model. Secondly,sparse representation theory is applied to low-quality fingerprint image denoising to enhance the texture structure of the fingerprint. Thirdly,in response to the allocation inequity issue caused by client heterogeneity,this paper proposes a client scheduling strategy based on reservoir sampling. Finally,experimental results on three real-world databases show that Fed-FR significantly outperforms local learning by 5.32% and federated average by 8.56%,approaching the accuracy of centralized learning. The results demonstrate the effectiveness of Fed-FR in privacy protection,accuracy evaluation,and scalability. This study demonstrates for the first time the feasibility of combining federated learning with fingerprint recognition,enhancing the security and scalability of fingerprint recognition algorithms,and providing a reference for the application of federated learning in biometric technologies.

Keywords　Fingerprint recognition,Federated learning,Sparse representation,Reservoir sampling,Privacy protection

1　引言

随着智能手机和其他设备上的生物识别扫描仪越来越普遍,以及越来越多的服务要求高安全性和良好的客户体验,生物特征识别技术逐渐淘汰口令、智能卡等传统手段,占据身份认证系统中的核心位置。在诸多生物特征识别技术中,指纹识别因其独特性与可靠性而受到用户广泛青睐。指纹识别系统需要大量的用户主动参与训练任务,并希望用户能够上传

基金项目:国家重点研发计划(2023YFB4302901,2023YFB2703700);国家自然科学基金(61802276,62172418,U2133205,U21A20465);天津市教委科研计划项目(2021KJ038);浙江理工大学科学基金项目(22222266Y)
This work was supported by the National Key Research and Development Program of China(2023YFB4302901,2023YFB2703700),National Natural Science Foundation of China(61802276,62172418,U2133205,U21A20465),Scientific Research Project of Tianjin Educational Committee(2021KJ038) and Science Foundation of Zhejiang Sci-Tech University(ZSTU)(22222266Y).
通信作者:鲁艳蓉(yr_lu@cauc.edu.cn)

高质量指纹数据，为公众提供更好的服务质量[1]。近年来，机器学习中自动学习与调整算法的使用，提高了识别效率和准确度。然而，由于设备和环境等外部因素的复杂组合，采集和传输的过程中会不可避免地受到外界噪声的干扰，从而降低了系统识别的准确性。此外，用户不愿意上传隐私数据。一方面，指纹的特征有可能会威胁到个人隐私，一旦指纹被攻击、复制和泄漏，会导致指纹永久丢失。另一方面，如果一个人的指纹存储在多个系统中，那么攻击者可能通过交叉映射获得。因此，研究在隐私约束的前提下模型准确率与公平性训练机制，对进一步扩展指纹识别应用规模具有重要意义。

现有解决方案利用机器学习算法对指纹图像集中处理完成模型训练。Zhang 等[2]使用 KNN 算法改进 Triplet-Loss 函数的样本选择问题，并使用 SoftMax Loss 进行卷积神经网络训练，有效提高网络表达指纹图像的能力。Nogueira 等[3]使用卷积神经网络进行指纹活体检测，并准确地检测出伪造或篡改的指纹。尽管这些方法取得了较高的准确性，但这些算法本身存在安全漏洞且并未对数据进行隐私保护，直接使用可能会给系统带来潜在的威胁。此外，机器学习的训练集和测试集样本来自于同一独立数据分布[4]。当数据来自不同的分布时，分类器将失去判断，这需要模型通过扩大用户数量来学习更多数据分布特征。Zhao 等[5]提出在边缘设备之间共享一小部分数据，以改善非独立同分布数据对的模型泛化能力。然而，由于数据隐私的限制，这种策略是不可行的。随着处理器的快速发展，终端设备成为在数据源上执行机器学习的理想平台，可为每个用户提供一个安全的平台，并按照需求进行训练。然而，由于上传的数据和接收更新后的模型需要消耗大量带宽，从而使得通信资源受限，每个用户需要花费大量的时间成本来保障。因此，传统集中式的方法受隐私保护限制，不能扩展于大量用户的应用背景。

联邦学习[6](FL)是当前人工智能背景下实现数据隐私保护，利用分散在各参与方的数据集，通过隐私保护技术融合多方数据信息，协同构建全局模型的一种分布式训练方法。每一个数据源客户端利用自身的数据单独训练一个模型，之后各终端的模型彼此之间进行交互，避免在学习过程中泄露用户的原始数据。联邦学习平均算法[7](Federated averaging，FedAvg)在服务器上聚合并平均本地客户端的梯度，并将更新后的模型传回客户端进行下一轮本地优化。近年来，联邦学习在图像分类任务上取得显著进展，提高了聚合全局模型在多种联邦学习场景下的分类性能。然而，由于以下两个关键原因，联邦学习无法被直接应用于指纹识别系统中。

(1)图像质量的多样性会导致模型聚合精度降低[8]。指纹自动识别的一个主要问题是原始指纹的质量，如果质量达不到可接受的标准，将无法进行特征提取。联邦学习系统由大量本地客户端组成，而这些用户数据可能包含噪声，这将严重阻碍全局模型收敛效果。

(2)异构数据容易产生联邦学习的不公平性。由于通信资源的限制，只有一小部分客户端可被挑选出来代表模型所有者进行训练[6]。不同的客户端有自己的数据分布特征。当数据是非独立同分布时，不同客户端的梯度之间的差异会变得更大。因此，优先级低的客户端容易被剥夺参与的机会，产生客户端之间的不公平选择。这种极端的选择方案，无法保证数据的多样性和全面性，从而在一定程度上损害了模型的性能。

有别于传统集中训练的指纹识别，为实现在隐私保护下提高模型聚合的准确率，本文提出了基于联邦学习的指纹识别算法(Fed-FR)，以隐私感知的方式进行指纹识别模型训练。此外，为解决客户端数据异构性问题，提出了一种基于水库抽样[9]的客户端选择策略，进而实现客户端选择的公平性。主要贡献如下：

(1)为满足用户对数据隐私保护的需求，提出了联邦学习指纹识别算法 Fed-FR。该算法可在严格遵守隐私约束的情况下，有效地改进指纹识别模型。

(2)为降低样本质量对全局模型的影响和实现客户端选择的公平性，结合稀疏表示理论及水库抽样算法，设计了一个样本去噪模块与客户端选择策略，从而提升算法识别率与确保整体网络服务质量的公平统一。

(3)实验结果表明，与无联邦学习的指纹识别模型相比，Fed-FR 指纹识别模型完成隐私保护的同时，在识别过程中获得了较高的识别率且在客户端数量上具有可扩展性。

本文第 2 章简要介绍指纹识别数据安全与准确性相关工作的研究现状；第 3 章介绍相关概念及问题描述；第 4 章介绍基于联邦学习的指纹识别算法设计；第 5 章对所提方法进行仿真实验；第 6 节进行性能分析；最后总结全文并展望未来。

2 相关工作

2.1 指纹识别

目前，最先进的基于三维数据结构的圆柱编码(Minutia Cylinder Code，MCC)[10]指纹匹配算法在数据库 FVC2002 DB2 上的 $EER=0.49\%$，在数据库 FVC2006 DB2 上的 $EER=0.12\%$。Lee 等提出一种新的部分指纹匹配方法[11]，通过对特定边缘形状的小脊线段进行特征提取，使得在脊线特征相似度计算阶段，只包含相似指纹重叠部分的特征，提高了指纹识别的准确率。但是，生物识别系统中，由于生物特征不能被撤销或重置，如果原始的、不受保护的数据直接存储在数据库中，可能会出现严重的安全问题，如攻击者可侵入数据库中的模板数据，从而获得对生物识别系统的未经授权的访问。Alam 等[12]提出了一种生物密码系统的无对准可取消指纹模板，在训练过程中引入噪声，在不对对手的背景知识建模的情况下，隐藏个人隐私信息，同时保持整个聚合数据集的基本统计特征，因此，攻击方无法检测用户的存在或不存在。Wang 等[13]利用低秩矩阵近似来降低指纹的维数确定最优秩，并将拉普拉斯噪声注入到近似奇异矩阵的奇异值中完成指纹识别。但是，差分隐私存在隐私和效用的冲突，难以提供严格的安全保证，且要付出准确性损失。

随着深度神经网络的广泛应用，研究者发现利用神经网络强大的特征提取能力能够实现良好的指纹识别[14]。Cappelli 等[15]提出了一种基于线性支持向量机的指纹图像分割方法，将指纹图像的子块中提取的傅里叶谱能量比和灰度对比度构造成特征向量，并使用线性支持向量机对其进行分类，并进行形态学运算，从而实现指纹图像分割。Zhang 等[2]使用 KNN 算法改进三元组损失函数的样本选择问题，并使用 SoftMax 损失进行卷积神经网络训练，有效提高了网络表达指纹图像的能力。然而，无论是现有的基于传统手工特征的部分指纹识别算法，还是深度学习指纹识别算法，都存在一定的局限性。此外，由于用户的偏好不同，导致不同用户的数据呈现出不同的分布特征，从而出现标签分布偏差、特征分布偏差和权重差异等情况，这些情况将导致模型精度

降低、收敛速度变慢等问题。

2.2 联邦学习

传统密钥加密技术受限于密钥长度和加密算法本身的复杂度，难以有效抵御日益增强的计算能力和潜在的量子计算攻击，其安全性无法得到充分保障。在无线和异构网络中，密钥的分发、维护和管理面临诸多挑战，无法满足现代网络环境的动态性和复杂性需求。联邦学习允许多个实体在没有数据共享的情况下共同构建一个复杂、健壮的模型，从而解决诸如数据隐私、安全、访问权限以及机器学习分析中对异构数据的访问等关键问题[16]。联邦平均算法是联邦学习中经典的算法，通过聚合并平均服务器上本地客户端的梯度，并将更新后的模型传回客户端进行下一轮本地优化。然而，在每一轮通信中，服务器仅随机选择一个客户端子集[17]来使用其本地数据集训练全局模型，不能保证每一轮的数据分布代表真实分布。当服务器重复选择特定类型的客户端时，数据分布可能会产生高度不平衡现象。文献[18]提出了 Favor 方法。该方法通过智能地选择客户端设备参与每一轮联邦学习，可有效缓解非独立同分布数据带来的偏差，并加快模型收敛速度。

综上所述，现有的研究大多集中在指纹识别的隐私保护或性能提升的单方面问题，对在指纹识别系统中同时兼顾隐私保护、识别精度和扩展性问题的研究还比较少。因此，本文提出隐私约束下指纹识别方案，使得用户在可信环境下完成身份识别，并增加数据预处理模块与改进客户端选择策略，从而确保整体方案的准确性和公平性。

3 相关概念及问题描述

将联邦学习作为可信的模型训练平台，完成全局模型的参数聚合和更新，并通过稀疏表示理论实现本地数据的预处理，得到高质量和完整的指纹纹理结构。为保证客户端分配的公平性，通过水库抽样算法将所有客户端进行标记产生等同概率的随机子集，降低了客户端之间的冲突。

3.1 基于联邦学习的指纹识别

联邦学习是一种分布式机器学习技术，允许多参与方通过本地训练集按照指定算法（聚合策略）协同构建全局模型。在模型训练过程中，模型的相关信息（如模型参数、模型结构、参数梯度等）能够在各参与方之间交换（交换方式可是明文、数据加密、添加噪声等），但本地训练数据不会离开本地，如图 1 所示。这一交换不会暴露本地的用户数据，降低数据泄漏的风险。联邦学习模型可在各数据参与方之间共享和部署使用。

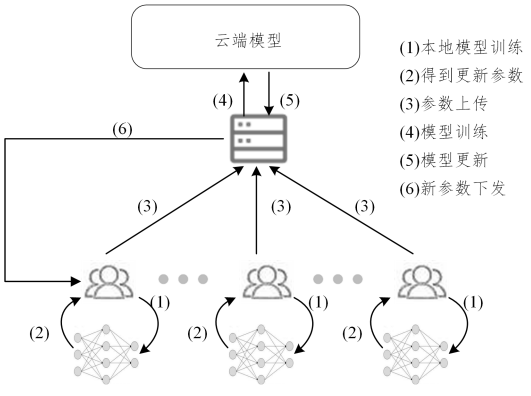


图 1 联邦学习架构

Fig. 1 Federated learning architecture

在联邦设置中，训练和测试中的类是不同的。在训练阶段，模型将输入指纹嵌入到高维表示中，并通过计算输入特征与所有类之间的相似性来生成损失值更新网络。

假设服务器在第一轮初始化全局模型，且客户端利用本地数据集更新优化器。通过实验固定学习率 η，在第 t 轮更新全局模型参数时，第 k 个参与方将会计算 $g_k=\nabla F_k(w_t)$，即它在当前模型参数 w_t 的本地数据的平均梯度。服务器会根据以下公式聚合全局梯度，并使用式(1)更新模型参数的信息。

$$w_{t+1}\leftarrow w_t-\eta\sum_{k=1}^{K}\frac{n_k}{n}g_k \tag{1}$$

其中，$\sum_{k=1}^{K}\frac{n_k}{n}g_k=\nabla f(w_t)$。接收客户端上传的模型，使用聚合函数更新全局模型，将更新后的模型参数 w_{t+1} 发送给各参与方。每个客户端 i 将在本地数据集上稀疏分解图像去噪，上传模型。

3.2 稀疏表示

稀疏表示已被广泛应用在众多领域，比如指纹图像压缩、指纹匹配、人脸识别、机器视觉和模式识别等。在指纹图像去噪中，通过基本元素的组合表达原图像 x 的过程中自动舍弃冗杂的信息 s，如式(2)所示：

$$y=x+s \quad \text{s.t.} \ \| y-x \|_2^2\leqslant\sigma^2 \tag{2}$$

其中，y 为自然采集数据，σ 为噪声标准方差。

指纹识别是一种通过解析响应的信息来判断用户身份的技术。然而，低质量、高存储是指纹识别技术面临的挑战。基于稀疏表示理论指出可从采样数据中去噪并重构指纹图像，从而增强图像质量。其理论在于指纹图像的稀疏性，相比于非稀疏的自然图像，指纹图像通常包含大量重复的脊和谷，这就为指纹图像稀疏性提供了前提条件。因此，将指纹图像稀疏化表示，只有少数系数较大，其它系数近似等于零，如图 2 所示。

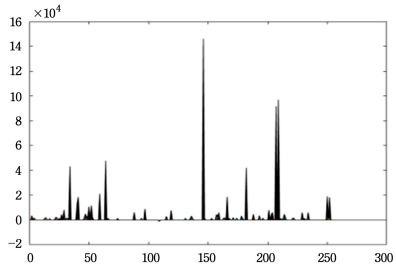


图 2 指纹图像的稀疏分解

Fig. 2 Sparse decomposition of fingerprint images

指纹图像是可稀疏的，即可通过有效个原子来表示，而噪声是不可稀疏的，可通过 y 去提取 x 的稀疏系数，再通过稀疏部分来重构图像。在这个过程中，噪声被处理为观测图像和重构图像之间的残差，重构过程中残差被丢弃，以达到去噪效果。将图像分块处理为大小为 $\sqrt{n}\times\sqrt{n}$ 的局部块 q，并且可通过有效误差进行稀疏表示。对于其给定的字典 A，可构造稀疏表示模型：

$$\hat{\alpha}=\arg\min_{\boldsymbol{\alpha}}\| \boldsymbol{\alpha} \|_1 \quad \text{s.t.} \ \| A\boldsymbol{\alpha}-q \|_2^2\leqslant\varepsilon \tag{3}$$

其中，$\boldsymbol{\alpha}$ 为 x 的稀疏系数向量，$\varepsilon=cn\sigma^2$ 且 c 为常数。

3.3 水库抽样

水库取样算法是经典的取样方法，当河流长度不确定或无限长时，对河流进行单次采样而不进行替换。假设所需样

本的大小为 k，算法通过保留流的前 k 项，然后以 $f(k,n)=k/n$ 的概率对每个后续元素进行采样。通过算法 1 可看出，在任何时候，流中的所有项目都以相等的概率出现在水库中[9]。它能够保证储层中的样本是均匀随机的样本，即使在流的演化过程中也没有意外的偏差，且仅需要 O(1) 时间（独立于水库大小和流的长度）处理流中的每个项目。因此，对于接收到的大于 k 的 n 个数据，确保选择的客户端为均匀采样。

算法 1 水库抽样算法

输入：最大储存大小 k，所有客户端数量 n

输出：等同概率选中客户端

1. 初始化容器
2. n:=1
3. for 流中的每一项 do
4. if n<k then
5. 将当前项插入储存器中
6. else
7. 以 f(n,k) 的概率，将随机均匀选择的储层元素弹出，并将当前项插入储层
8. end if
9. n:=n+1
10. end for

3.4 问题描述

通常在服务器上利用公开的数据集进行集中训练来部署模型，为使模型可适应新的用户加入时带来新的数据分布，需要收集存储在本地设备中的图像，用新数据更新模型。然而，由于用户隐私问题，禁止上传任何与身份相关的信息，如指纹原始图像及其特征。联邦学习提供一个框架来训练模型，在该框架中，假设有 K 个参与方，设 D_k 表示由第 k 个参与方所拥有的数据集，P_k 表示位于客户 k 的数据点的索引集，设 $n_k=|P_k|$ 表示 P_k 的集合大小。因此，在第 k 个参与方有 n_k 个数据点时，K 个参与方损失函数（如交叉熵损失）为[7]：

$$f(w)=\sum_{k=1}^{K}\frac{n_k}{n}F_k(w),F_k(w)=\frac{1}{n_k}\sum_{i\in P_k}f_i(w) \tag{4}$$

然而，当应用于大规模指纹库时无法避免两方面问题。一方面，本地客户端之间采集环境不同，易产生低质量数据。由于联邦学习系统由众多客户端组成，这些用户数据可能包含有错误或噪声样本，从而阻碍全局模型收敛。为能有效地去除图像噪声，同时又能最大限度地保留指纹的纹理结构，利用 K-SVD 算法对离散余弦变换字典进行自适应更新，以应对低信噪比场景下的噪声去除问题。另一方面，由于客户端用户的偏好不同，导致不同客户端上的数据呈现出不同的分布特征，从而出现标签分布偏差、特征分布偏差、权重差异等情况[6]。每当全局参数与局部参数发生冲突时，第 t 轮未被选择的客户端可能会存在被模型遗忘的风险。为解决这个问题，利用算法 1 将客户端选择问题转化为多次循环抽样问题，从而所有参与的客户端共同构建公平的和等同概率选中的随机子集。Fed-FR 旨在解决以上问题。

4 算法设计

针对指纹识别背景下联邦学习中存在的低质量原数据、客户端公平性选择等问题，基于稀疏表示理论与水库抽样算法，提出了一种新的指纹识别算法 Fed-FR。

4.1 系统架构

本文提出了一种新的基于联邦学习的指纹识别算法 Fed-FR。Fed-FR 的整体系统架构如图 3 所示。通过联邦学习将多个本地模型结合起来训练一个通用识别器，而不破坏客户端的用户隐私。

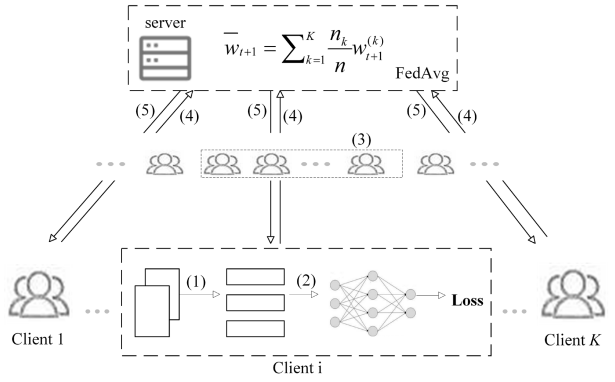


图 3 算法架构

Fig. 3 Algorithmic architecture

为得到高质量的指纹图像，本文引入稀疏表示理论及方法，得到去噪后的完整的指纹信息；为解决联邦学习设置中客户端之间更公平的准确性分布的问题，提出了一种客户端公平调度策略。整体算法流程如图 4 所示。当参与训练集合中新客户端到达时，只需要决定是否插入到当前样本数组中，然后在集合随机位置覆盖样本数据。

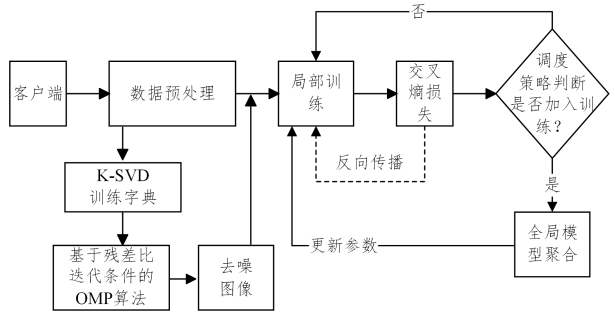


图 4 算法流程

Fig. 4 Algorithmic processes

客户端通过其本地数据以联邦学习方式训练识别模型，以构建更健壮的全局模型。框架中，采用横向联邦学习系统的客户端-服务器体系结构来实现所涉及的算法，中心服务器负责将初始模型发送给各参与者 $\{1,2,\cdots,K\}$。其中，K 表示参与方的数量，通常 K 的值大于或等于 2。在第 t 轮通信中，服务器向所有客户端节点发送全局模型 w_t，其中数据拥有者 $\{1,2,\cdots,K\}$ 分别使用各自的本地数据集来更新初始模型，并将更新后的模型权重参数发送给聚合服务器。服务器将从数据拥有者处接收到模型参数并聚合，通过联邦服务器中 FedAvg 算法更新全局模型参数，服务器将聚合后的全局模型重新发送给各参与方，进行下一轮的联邦训练。这一过程重复进行。在新的迭代开始前，公平地选择每轮所参与的客户端，直至模型收敛。由此，可避免传统机器学习方法集中学习而产生的隐私性问题。

为处理低质量指纹图像，采用 DCT 冗余字典、K-SVD 算法和基于残差比迭代终止条件的 OMP 算法将加性噪声去除，从而有效地保留指纹的边缘结构和纹理细节，改善视觉效果和识别性能。此外，引入水库抽样算法从所有客户端中公平地选择子集，显著减轻整体训练过程偏向于部分客户端的

问题，确保数据分布的完整性。

因此，可以通过多个客户端公平性的协助学习，结合本地目标更新局部模型实现全局模型的最佳性能。

4.2 数据预处理模块

在客户端加入模型训练前完成本地数据的预处理，在各个客户端上使用低质量图像本身的补丁训练字典，利用稀疏表示进行图像去噪，以提高数据集质量。假设一个固定的 $\boldsymbol{D}$ 字典和 $\boldsymbol{X}$ 原始图像，通过部署 OMP 算法完成稀疏编码阶段。

首先，从原始样本 $\boldsymbol{Y}\in R^{m\times n}$ 中随机取 K 个列向量或者取它的左奇异矩阵的前 K 个列向量$\{\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_K\}$作为初始字典的原子，得到字典 $D^{(0)}\in R^{m\times K}$。其次，利用字典 $D^{(j)}$ 稀疏编码，得到对应的编码系数 $\boldsymbol{X}^{(j)}\in R^{K\times n}$。最后，在满足稀疏条件下逐列更新字典 $D^{(j)}$，通过字典的列 $\boldsymbol{d}_k\in\{\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_K\}$，找出使用原始的补丁集合 $w_l=\{(i,j)\,|\,\alpha_{ij}(l)\neq 0\}$，并利用其索引计算误差矩阵：

$$\boldsymbol{e}_{ij}^{l}=\boldsymbol{R}_{ij}\boldsymbol{X}_{ij}-\sum_{m\neq l}\boldsymbol{d}_m\alpha_{ij}(m) \tag{5}$$

其中，j 的初始值为 0，$\boldsymbol{X}_{ij}$ 为子图提取矩阵，$\boldsymbol{a}_{ij}$ 为稀疏矩阵。

应用 SVD 分解 $\boldsymbol{E}_l=\boldsymbol{U}\Delta\boldsymbol{V}^{\mathrm{T}}$，更新系数，直到达到指定的迭代步数。首先，使用 K-SVD 建立更新字典。在找到输出图像 $\boldsymbol{X}$ 之前，选择使用相同的值执行更多的表示计算和字典更新迭代。其次，使用式(5)得到输出去噪图像。

假设指纹图像用矩阵 $\boldsymbol{Y}\in R^{N}$ 表示，大小为 $N\times N$，字典矩阵 $\boldsymbol{D}$ 由 L 个 N 维向量的 $\boldsymbol{d}_{ij}$ 组成，$\boldsymbol{d}_{ij}$ 称为字典原子。当 $\boldsymbol{D}$ 给定时，图像 $\boldsymbol{Y}$ 可表示为字典原子的线性组合：

$$\boldsymbol{Y}=\sum_{i=1}^{L}a_{ij}d_{ij}+R^{(M)} \tag{6}$$

其中，a_{ij} 为 $\boldsymbol{Y}$ 在字典 $\boldsymbol{D}$ 上的稀疏表示系数；矩阵 $\boldsymbol{R}^{(M)}$ 是经过 M 项逼近后的残差项。采用 K-SVD 方法进行字典训练得到字典 $\boldsymbol{D}$，并将其用于图像去噪。用矩阵 $\boldsymbol{X}$ 表示原始图像，则图像的去噪模型为：

$$\min_{a_{ij},\boldsymbol{X},\boldsymbol{D}}\lambda\,\|\boldsymbol{X}-\boldsymbol{Y}\|_2^2+\sum_{ij}\mu_{ij}\,\|a_{ij}\|_0+\sum_{ij}\|\boldsymbol{D}a_{ij}-\boldsymbol{R}_{ij}\boldsymbol{X}\|_2^2 \tag{7}$$

其中，$\lambda\,\|\boldsymbol{X}-\boldsymbol{Y}\|_2^2$ 为似然项，表示观测图像 $\boldsymbol{Y}$ 与原始图像 $\boldsymbol{X}$ 的逼近程度；$\sum_{ij}\mu_{ij}\,\|a_{ij}\|_0$ 为稀疏性约束；$\sum_{ij}\|\boldsymbol{D}a_{ij}-\boldsymbol{R}_{ij}\boldsymbol{X}\|_2^2$ 表示重建子图与原子图的相似性；$\boldsymbol{D}a_{ij}$ 为重建子图矩阵；$\boldsymbol{R}_{ij}$ 是子图提取矩阵。

初始化字典 $\boldsymbol{D}$ 和图像 $\boldsymbol{X}$，利用正交匹配追踪算法求得稀疏表示系数 a_{ij}；其次，根据得到的稀疏系数 a_{ij}，利用 K-SVD 算法更新字典 $\boldsymbol{D}$；最后，当 $\boldsymbol{D}$ 和系数 a_{ij} 均达到要求时，根据下式求得去噪图像：

$$\hat{\boldsymbol{X}}=(\lambda\mathbf{I}+\sum_{ij}\boldsymbol{R}_{ij}^{\mathrm{T}}\boldsymbol{R}_{ij})^{-1}(\lambda Y+\sum_{ij}\boldsymbol{R}_{ij}^{\mathrm{T}}\boldsymbol{D}a_{ij}) \tag{8}$$

其中，$\mathbf{I}$ 为单位矩阵，$\hat{\boldsymbol{X}}$ 为去噪后的图像矩阵。

4.3 客户端调度策略

客户端调度策略的整体方案如图 5 所示。当参与训练集合中新客户端到达时，只需要决定是否插入到当前样本数组中，随后在集合随机位置覆盖样本数据。以往每一轮通信中，服务器仅随机选择一个客户端子集来使用其本地数据集训练全局模型，不能保证每一轮的数据分布代表真实的数据分布。当服务器重复选择特定类型的客户端时，数据分布可能会高度不平衡，梯度的平均值可能会被重复选择的客户端所支配[18]。为获得更公平的权衡分配，本文提出了一种客户端选择策略，从数据流中无偏挑选样本。在第 i 个本地客户端上的第 t 轮迭代轮次开始时，首先，从数据流中维护一个大小为 n 的容器。数据流中的前 n 个客户端被添加到储存中进行初始选择。其次，当接收到数据流中第 $n+1$ 个客户端时，以 $n/k+1$ 的概率添加到容器中，这个客户端会取代容器中随机选择的客户端。最后，当流中新客户端节点到达时，只需决定是否插入到子集的当前样本数组中，可在随机位置覆盖样本。

4.4 本地模型训练

指纹识别模型需要使用广泛的数据进行训练，以获得不同的数据属性，从而正确地对用户进行识别。给定客户端 i 上的初始模型参数，本地通过交叉熵函数[19]对所设计的网络进行训练，将输出的节点的值限制在[−1,1]，避免后续特征计算过程中的特征向量归一化操作，如图 4 所示。同时，利用梯度下降来求解并更新参数。

在联邦设置中，随机分配训练数据，使得每个客户端只有特定用户指纹的数据样本，以确保数据的异质性。每个客户端的模型都是根据其私有数据进行本地训练，周期性地执行全局模型聚合。加入联邦训练后，由于客户端之间梯度不进行共享，因此，在服务器上完成梯度聚合。在服务器端，随着迭代次数增加，利用算法 1 删除已参与的客户端，加入未参与的新节点。该算法需要一定的内存来保持样本集中每个客户端的索引信息，保证公平地学习到完整数据分布。此外，网络输出的规模不等于客户端的数量，可在不需要知道网络参与者数量的情况下，解决客户端数量的可扩展性的问题。在训练步骤中，本地模型结合本地数据集进行训练。首先，完成 FedAvg 和模型更新。FedAvg 是联邦随机梯度下降的泛化，允许所有客户端批量更新数据并交换更新的权重而不是梯度。因为联邦随机梯度下降中的大多数局部客户端都是从一个共同的初始化开始的，因此，平均梯度等价于平均权重。其次，从相同的初始化中选择调优权重并不一定会影响平均模型的性能。最后，经过训练的模型可执行测试集的指纹识别过程。

综上所述，算法 2 提出指纹识别中的联邦学习架构，该模型可让多客户端协同合作且有效地利用本地数据更新全局模型，保护用户隐私，实现高性能的指纹识别。首先，模型在整体算法前进行初始化。在该步骤中，不仅联邦服务器获得初始模型，同时初始化了全局模型和局部模型的权重、梯度和超参数。其次，进入客户端本地更新步骤前，需要在本地客户端上完成数据的预处理，增强指纹图像的纹理结构。最后，通过算法 1 公平地选择客户端参与训练，直至模型收敛，获得全局模型。

算法 2 联邦学习指纹识别算法

输入：终端设备集合 N，批处理 B，初始化参数 w_0，每轮参与训练客户端数 k，学习率 η

输出：全局模型参数

服务器端：

1. 初始化模型参数
2. for 全局更新轮次 t=1,2,…,do
3. 通过算法 1 选择客户端子集 max(k,1)
4. for 每一参与方 k do
5. $w_{t+1}^{(k)}\leftarrow$参与方更新(k,$\bar{w}_t$)
6. 服务端判断模型是否收敛
7. end for

8. 服务端加权平均：$\bar{w}_{t+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} w_{t+1}^{(k)}$

9. 客户端更新(k，$\bar{w}_t$)

10. for 本地迭代次数的每一次本地迭代 i do

11. $w_{1,i}^{(k)} = w_{B,i-1}^{(k)}$

12. 本地更新参数：$w_{b+1}^{(k)} \leftarrow w_{b,i}^{(k)} - \eta g_k^{(b)}$

13. end for

14. end for

15. 模型收敛，获得全局模型

5 实验

5.1 实验数据集及环境

为使实验具有可比性，本文选择主流的 3 个公共数据集：Sokoto Coventry Fingerprint Dataset[20]，SFinGe[21]，FVC 2004[22]。

(1)SOCOFing 是一个专为学术研究而设计的数据库，包含独特的属性，如性别、手和手指名的标签。数据库中有 6000 枚真实指纹和 17931 枚合成变造指纹图像。每幅图像的大小为 96×103 像素。全局模型训练中选取 6000 幅指纹图像进行测试。图 6 为原始图像中 1 个受试者采集的样本。每个受试者采集 10 张指纹图像，采集过程中对指纹进行 z 切割、消除和中心旋转等处理。低质量指纹与清晰指纹的比例设置为 3∶2。从每个人的指纹中随机抽取 6 张指纹图像作为训练集，其余指纹图像作为测试集。

图 6 同一受试者的十枚指纹

Fig. 6 Ten fingerprints from the same subject

(2)SFinGe 是一个合成数据库，其中包含 50 个低质量指纹，这些指纹在不同压力和干燥度水平下生成，并且每个手指都有两个印记。同时，通过不同程度的旋转、不同距离的平移，裁剪和转换等方式，共生成了 300 枚指纹。

(3)FVC 2004 数据库有 4 组(DB1，DB2，DB3 和 DB4)灰度图像，每组有 80 个指纹，且每个手指产生 8 张图像。数据库中共有 640 张指纹图像。其中，DB1－DB3 包含真实指纹，而 DB4 包含合成指纹。

指纹识别系统的性能主要通过区分真假指纹的能力来衡量。识别率为系统正确识别的指纹数与数据库中指纹总数的比值，是衡量指纹识别系统整体性能的常用指标。本文在 Intel(R)Core(TM)i3-9100F CPU 处理器和 RTX 3090 的 GPU 处理器上仿真模拟联邦学习环境，实验程序部署于 Anaconda 和 PyCharm 联合开发环境，核心库版本 python＝3.7，pytorch＝1.1。

5.2 实验参数选择

本节旨在通过仿真实验选择合适参数。基于稀疏表示的指纹图像去噪，需要选择合适的补丁大小。补丁的大小直接影响着去噪效果。然而，到目前为止，还没有很好的方法来估计这个参数，一般通过实验进行选择。考虑 8×8，12×12 和 16×16 的大小，性能结果如图 7 所示。8×8 这种尺寸的补丁太小，无法完整包含指纹的结构。16×16 大小在高压缩下会产生较高的复杂度。如图 7 所示，面对低质量指纹图像时，大小为 12×12 效果最好。

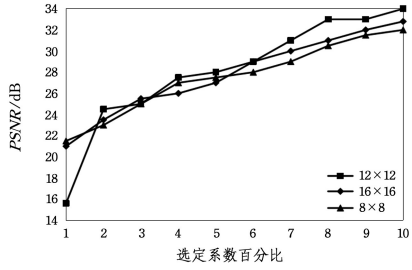


图 7 不同补丁的性能结果

Fig. 7 Performance results of different patches

在设计框架时考虑参数值会影响模型的收敛速度。因此，在联邦训练中，对不同的参数进行实验分析，寻求超参数的最佳组合。如局部迭代次数和批处理大小，并在后文中使用统一的超参数进行对比验证。设置局部迭代轮次 $E=\{1, 5\}$，局部批处理大小 $B=\{10, 16, 50\}$，在数据集上使用相同的学习率 $\eta=0.01$ 进行实验。

横坐标为全局迭代轮次，纵坐标为模型识别精度，如图 8 所示。当 $B=16$，$E=5$ 时，在数据集上取得了最高精度。实验结果表明，本地迭代轮次数过少会导致模型训练不完全，而本地迭代轮次过多会增大不同客户端之间的模型差异，且随着局部批处理大小的增加，模型收敛速度下降。因此，需要在实验过程中设定固定的参数进行实验。

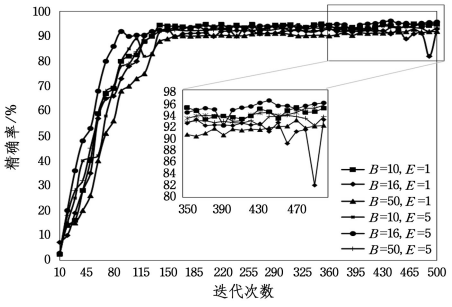


图 8 不同超参数下的模型性能比较

Fig. 8 Performance comparison of models with different hyperparameters

5.3 不同模型的测试结果

本节旨在探讨不同模型对算法性能的影响。通过比较 SOCOFing 数据集上不同模型的识别精度来选择合适模型。实验中，使用 AlexNet，GoogleNet，ResNet18 这 3 个模型进行验证。如图 9 所示的曲线变化表明，随着训练次数的增加，模型的准确性逐渐增加，而后趋于稳定。

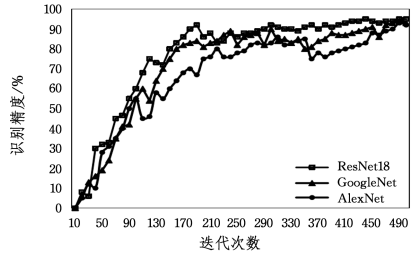


图 9 不同模型下的识别性能

Fig. 9 Performance of recognition of different models

实验过程中，全局模型迭代 200 轮次时，AlexNet 模型训练精度明显低于其余模型，在相同超参数时性能并不理想；而 ResNet18 在同样配置下具有更稳定的精度表现，且相比其余模型在较短的轮次内达到了 95% 的精度，完成了模型收敛。本地客户端采用 ResNet18 模型完成本地模型训练。

5.4 不同去噪算法的测试结果

本节旨在探讨不同的去噪方式对算法性能的影响，并通过比较 SOCOFing 数据集上不同的指纹和噪声水平下去噪的峰值信噪比（Peak Signal to Noise Ratio，PSNR）数据来选择合适方法。使用基于 DCT 过完备原子库 OMP 去噪算法[23]、Contourlet 去噪算法和基于 K-SVD 残差低信噪比去噪算法进行对比实验。实验中，使用的字典大小为 64×256，设计用于处理 12×12 像素的图像补丁。本地客户端指纹图像的去噪结果如表 1 所列。

表 1 PSNR 数据比较

Table 1 Comparison of PSNR data

标准差	峰值信噪比/dB		
	Contourlet	DCT	K-SVD
2	38.43(38.47)	39.10(39.53)	43.31(43.05)
5	37.49(37.65)	37.51(37.79)	39.62(39.95)
10	36.48(35.61)	36.28(36.42)	36.50(37.05)
15	34.61(34.13)	33.38(33.93)	35.70(35.89)
20	32.60(32.31)	31.86(31.96)	33.74(33.40)
25	31.27(30.17)	30.95(39.71)	31.73(31.82)
50	28.69(29.02)	29.89(30.42)	30.92(30.57)
75	28.81(25.45)	24.16(23.12)	25.41(26.02)
100	24.64(24.26)	24.45(24.95)	25.03(25.36)

表中括号内为每个结果重复 5 次实验的平均值。由表 1 可看出，在强噪声环境下，基于 Contourlet 变换的去噪算法并不能较好地去噪。这是因为在去噪过程中，算法可能将指纹的边缘部分与高频信息当作噪声而进行过滤。基于 DCT 过完备原子库 OMP 去噪算法易在重构时误将部分噪声认为有用信号而影响重构精度，而 K-SVD 方法学习到的字典重构误差更小，可充分重构指纹的纹理结构，对于客户端，加入全局训练能有效提高识别精度。

6 性能分析

6.1 Fed-FR 在客户端上的公平性

本节旨在探讨 Fed-FR 在客户端公平性上的优势。在仿真实验中，将参与联邦学习的客户端总数量设置为 100，每个用户分配 600 张训练图片。为模拟数据独立同分布和非独立同分布情况，对均匀分布的数据集进行随机均匀分配以满足独立同分布，对每个参与客户端只分配部分数据以满足非独立同分布。

实验中，通过调整独立同分布客户端在总客户端所占比例（0.3，0.6，0.9 和 1）完成对比。为验证 Fed-FR 的公平性，使用 FedAvg 和 FedProx[24] 作为对比算法。如表 2 所列，在绝对理想情况下，即客户端为独立同分布设置时（设置为 1 时），Fed-FR，FedAvg 和 FedProx 这 3 种算法对数据的识别精度差别很小，其原因为客户端上的权值差异基本相同。在异构客户端时，存在非独立同分布客户端。在设置为 0.3 时，Fed-FR 算法识别的准确率高于对比算法，相比 FedAvg 算法提高了 13.279%。可以看出，FedAvg 在理想情况下可以达到较高识别率，但当出现异构客户端时，识别的性能是不稳定的。在设置为 0.6 时，FedProx 算法也出现识别不稳定的情况。

表 2 不同算法在测试集上的准确率

Table 2 Accuracy of different algorithms on test set

比例	FedAvg/%	FedProx/%	Fed-FR/%
0.3	81.74	89.62	95.02
0.6	71.36	80.49	92.62
0.9	53.29	66.86	80.92
1	92.11	94.38	96.32

6.2 Fed-FR 在真实指纹上的性能分析

本节旨在探讨 Fed-FR 在真实指纹上的识别性能的有效性。如表 3 所列，Fed-FR 算法在 SOCOFing 上的识别率为 95.02%，在 SFinGe 上为 94.05%，在 FVC 2004 上为 94.82%，都具有较高的精度，验证了 Fed-FR 算法在不同数据集上的可迁移性，同时面对失真指纹具有较强的稳定性。

表 3 Fed-FR 算法精确度

Table 3 Accuracy of Fed-FR algorithm

数据集	识别率/%	识别出指纹数量	未识别出指纹数量
SOCOFing	95.02	5702	298
SFinGe	94.05	282	18
FVC2004	94.82	606	34

将 Fed-FR 与最近的指纹识别方法进行对比，Fed-FR 具有相对更高的识别精度，同时也为客户端提供更多安全保障，如表 4 所列。具体而言，文献[25]与文献[26]的识别精度略高于 Fed-FR，但 Fed-FR 则在保障精度的同时完成了用户的隐私保护。Fed-FR 在识别性能上优于文献[27]和文献[28]，这是由于 Fed-FR 将多个客户端的数据进行学习，对多元的数据分布学习得更加充分，从而降低了由客户端异构导致的分配不公的影响，提高了模型性能。从识别结果中可看出，Fed-FR 的识别精度仅略低于集中学习，并有效地保护了本地客户端的数据隐私。

表 4 比较结果

Table 4 Comparison results

算法	识别准确度/%
CNN[25]	95.50
DRB+DBM[26]	95.16
ANN[27]	86.82
BPN[28]	94.90
Fed-FR	95.02

6.3 消融实验

本节旨在探讨数据预处理和客户端选择策略对算法性能的影响。实验基于 ResNet 模型，在 SOCOFing 数据集上进行。当不使用数据预处理模块时，本地客户端仅进行简单的数据清洗；当不使用客户端选择策略时，每轮训练中只随机选择一小部分客户端参与。实验结果如表 5 所列。

表 5 模块比较

Table 5 Module comparison

数据预处理	客户端调度策略	准确率/%
×	×	80.64
√	×	88.32
×	√	82.64
√	√	95.02

当仅使用数据预处理模块时，全局模型的准确率提高 7.68%；进一步增加客户端调度策略后，由于全局模型所学习到的数据分布更加均衡多样，使得全局模型的识别准确率达到 95.02%。实验结果表明，数据预处理和客户端调度策略能有效地缓解数据异质性与客户端异构性问题，从而进一步提高识别的准确率。

通过比较表6,从同一数据集上集中与联邦学习方式、数据预处理和客户端选择策略3个方面来评估系统的识别准确度。

基线方法:

(1)集中学习,设备上的所有本地数据都被收集到中央服务器进行训练;

(2)局部学习,随机挑选部分客户端;

(3)联邦平均学习。

本文通过消融实验来验证每个模块的有效性。在联邦设置下,全局模型用于通用评估,局部模型用于评估本地模型,其中显示的精度是所有本地客户端的平均精度。表中第一行是用360个本地用户数据集中训练模型,第2、3行是在隐私约束下使用240个用户数据加入训练,最后一行是在理想用户数量下集中优化数据模型。从表6可看出,在第二行中,直接使用本地数据优化模型,由于存在低质量指纹图像,难以判断缺失部分是否存在特征点,降低了算法对指纹的识别精度。可利用大量公共数据加入训练,但它可能会受到长时间训练和大量计算开销的影响。对在联邦学习框架中加入本地数据去噪的预处理模块,在隐私约束下,减少噪声部分的影响。在第三行中,性能得到明显提高。但是,在联邦学习公平性下,更一般化的选择将损坏识别客户端特定身份的性能。因此,在第4行中提出具有更公平性的Fed-FR架构,利用水库抽样算法保证客户端选择的公平性。如表6所列,使用去噪后的数据集与公平性客户端选择,二者的组合可提供更准确的特征并提高识别精度。

表6 消融实验

Table 6 Dissolution experiment

(%)

	数据预处理	客户端调度策略	全局模型精度	局部模型精度
集中学习(360IDs)	—	—	72.03	72.03
联邦学习(240IDs)	√	—	88.32	81.25
联邦学习(240IDs)	√	√	95.02	94.17
集中学习(600IDs)	√	—	98.58	98.58

6.4 Fed-FR客户端数量与可扩展性实验分析

本节旨在探讨客户端数量对算法性能的影响与可扩展性。图10展示了使用不同方法的实验结果。一般情况下,随着节点数量的增加,集中学习的平均准确率略有下降,这表明了测试集数据的异质性。

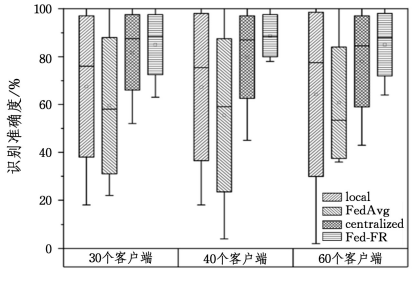


图10 节点数量对精度的影响

Fig. 10 Impact of node number on accuracy

由于FedAvg直接作用在数据集上无法收敛到中心化模型,导致FedAvg的效果最差。在同样配置下,Fed-FR的性能优于局部学习和联邦平均学习,并且精度超过40和60节点的集中学习,说明了Fed-FR的可扩展性。此外,集中学习的节点精确度的标准差随节点数量的增加变化较大,而Fed-FR的精度变化明显较小,这意味着Fed-FR可提高大多数节点的模型精度。

为验证Fed-FR的可扩展性,使用SOCOFing数据集的数据评估不同数量节点(30,40,60)的性能。如图11所示,模型的识别性能在测试集上是最好的,并且提出的模型在所有情况下都以非常低的假阳性率获得很高的真阳性率。因此,其能够将每个客户端更新的模型参数聚合到可信的聚合器上,以防止数据泄漏给模型所有者。在训练和验证集期间,通过增加客户端($N=30,40,50,60$)的数量,性能得到提高。该模型获得了较高的真阳性率,验证了模型随着客户端数量的增长而扩展,并且不需要客户端之间或客户端与联邦服务器之间的协调。

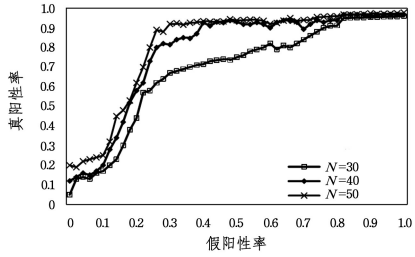


图11 ROC曲线

Fig. 11 ROC curve

6.5 单点实验分析

单点模型指的是在某一客户端 C_i,利用其本地数据 D_i 进行本地迭代训练的模型。为比较模型在识别阶段的性能,分别挑选其中的5个客户端来单独训练;联邦训练中分别设置不同的 k 值,表示每一次本地迭代训练,从所有客户端中挑选 k 个客户端来进行。实验中,分别设置 $k=3$ 和 $k=5$ 两个值。如图12所示,单点训练的模型效果明显低于联邦训练的模型效果,这表明仅通过单个客户端的数据不能很好地学习到数据的全部分布特性,模型的泛化能力较差。此外,每一轮参与联邦训练的客户端数目(k 值)不同,其性能也会有一定的差别,k 值越大,每一轮参与训练的客户端数目越多,性能越好,但每一轮的完成时间也会相对较长。

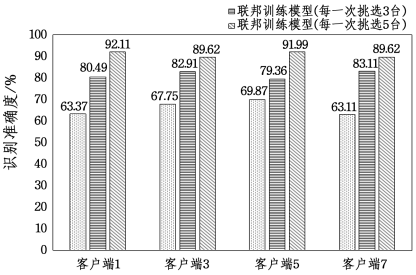


图12 算法流程

Fig. 12 Algorithmic processes

结束语 为同时兼顾指纹识别系统中用户的数据隐私和识别准确度,本文提出了一个联邦学习算法Fed-FR。该框架基于稀疏表示理论与水库抽样算法,采用K-SVD算法完成字典库的训练,利用OMP算法将加性噪声去除,从而得到高质量图像;依据客户端的公平性选择策略,对每个客户端进行标记使得其以相等概率参与聚合,从而完成模型完整数据分布学习。通过真实数据集完成实验评估,表明Fed-FR优于局部学习5.32%,FedAvg为8.56%,接近于集中学习的精度。深入研究Fed-FR在指纹识别任务中的适用性,与集中学习相比,其在效用上的损失可忽略不计。

虽然在这项工作中进行的相对小规模的实验验证了联邦学习对指纹识别任务的适用性，但是，传输的模型参数仍然揭示了用户信息存在的风险。在未来工作中，将结合密码学技术加强防御性能，实现更安全的隐私保护。

参 考 文 献

[1] HU Y，HU A，LI C，et al. Towards a privacy protection-capable noise fingerprinting for numerically aggregated data[J]. Computers & Security，2022，119(8)：102755.

[2] ZHANG F，FENG J. High-resolution mobile fingerprint matching via deep joint KNN-triplet embedding[C]// AAAI Conference on Artificial Intelligence. 2017：5019-2020.

[3] NOGUEIRA R F，DE ALENCAR LOTUFO R，MACHADO R C. Fingerprint liveness detection using convolutional neural networks[J]. IEEE Transactions on Information Forensics and Security，2016，11(6)：1206-1213.

[4] SUN S，CAO Z，ZHU H，et al. A survey of optimization methods from a machine learning perspective[J]. IEEE Transactions on Cybernetics，2019，50(8)：3668-3681.

[5] ZHAO Y，LI M，LAI L，et al. Federated learning with non-iid data[J]. arXiv：1806. 00582，2018.

[6] KAIROUZ P，MCMAHAN H B，AVENT B，et al. Advances and open problems in federated learning[J]. Foundations and Trends © in Machine Learning，2021，14(1/2)：1-210.

[7] MCMAHAN B，MOORE E，RAMAGE D，et al. Communication-efficient learning of deep networks from decentralized data[C]// Artificial Intelligence and Statistics. PMLR，2017：1273-1282.

[8] COETZEE L，BOTHA E C. Fingerprint recognition in low quality images[J]. Pattern recognition，1993，26(10)：1441-1460.

[9] LI K H. Reservoir-sampling algorithms of time complexity O(n(1+ log(N/n)))[J]. ACM Transactions on Mathematical Software，1994，20(4)：481-493.

[10] CAPPELLI R，FERRARA M，MALTONI D. Minutia cylinder-code：A new representation and matching technique for fingerprint recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2010，32(12)：2128-2141.

[11] LEE W，CHO S，CHOI H，et al. Partial fingerprint matching using minutiae and ridge shape features for small fingerprint scanners[J]. Expert Systems with Applications，2017，87(6)：183-198.

[12] ALAM B，JIN Z，YAP W S，et al. An alignment-free cancelable fingerprint template for bio-cryptosystems[J]. Journal of Network and Computer Applications，2018，115(4)：20-32.

[13] WANG T，ZHENG Z，BASHIR A K，et al. FinPrivacy：A privacy-preserving mechanism for fingerprint identification[J]. ACM Transactions on Internet Technology，2021，21(3)：1-15.

[14] ZHANG Z，LIU S，LIU M. A multi-task fully deep convolutional neural network for contactless fingerprint minutiae extraction[J]. Pattern Recognition，2021，120(12)：108189.

[15] CAPPELLI R，FERRARA M，MALTONI D. Minutia cylinder-code：A new representation and matching technique for fingerprint recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2010，32(12)：2128-2141.

[16] ZHAO J，CHANG X，FENG Y，et al. Participant selection for federated learning with heterogeneous data in intelligent transport system[J]. IEEE Transactions on Intelligent Transportation Systems，2022，24(1)：1-10.

[17] ZHAO J，CHANG X，FENG Y，et al. Participant selection for federated learning with heterogeneous data in intelligent transport system[J]. IEEE Transactions on Intelligent Transportation Systems，2022，24(1)：1-10.

[18] WANG H，KAPLAN Z，NIU D，et al. Optimizing federated learning on non-iid data with reinforcement learning[C]// IEEE INFOCOM 2020-IEEE Conference on Computer Communications. 2020：1698-1707.

[19] LI X，ZHANG X，HUANG W，et al. Truncation cross entropy loss for remote sensing image captioning[J]. IEEE Transactions on Geoscience and Remote Sensing，2020，59(6)：5246-5257.

[20] SHEHU Y I，RUIZ-GARCIA A，PALADE V，et al. Sokoto coventry fingerprint dataset[J]. arXiv：1807. 10609，2018.

[21] CAPPELLI R，MAIO D，MALTONI D. Synthetic fingerprint-database generation[C]// International Conference on Pattern Recognition. 2002：744-747.

[22] MAIO D，MALTONI D，CAPPELLI R，et al. FVC2004：Third fingerprint verification competition[C]// Biometric Authentication：First International Conference. 2004：1-7.

[23] DA CUNHA A L，ZHOU J，DO M N. The nonsubsampled contourlet transform：theory，design，and applications[J]. IEEE Transactions on Image Processing，2006，15(10)：3089-3101.

[24] LI T，SAHU A K，ZAHEER M，et al. Federated optimization in heterogeneous networks[C]// Machine Learning and Systems. MLSYS，2020：429-450.

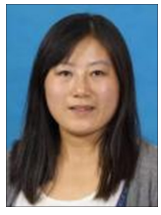
[25] NOGUEIRA R F，DE ALENCAR LOTUFO R，MACHADO R C. Fingerprint liveness detection using convolutional neural networks[J]. IEEE Transactions on Information Forensics and Security，2016，11(6)：1206-1213.

[26] ULIYAN D M，SADEGHI S，JALAB H A. Anti-spoofing method for fingerprint recognition using patch based deep learning machine[J]. Engineering Science and Technology an International Journal，2020，23(2)：264-273.

[27] PATIL S R，SURALKAR S R. Neural network based fingerprint classification[J]. International Journal of Science and Research，2013，2(1)：58-62.

[28] KOUAMO S，TANGHA C. Fingerprint recognition with artificial neural networks：application to e-learning[J]. Journal of Intelligent Learning Systems and Applications，2016，8(2)：39-49.

WANG Chenzhuo，born in 1999，postgraduate. Her main research interest is federated learning.



LU Yanrong，born in 1985，Ph.D，associate professor. Her main research interests include future cybersecurity，AI security and blockchain technology.