FedHome: Cloud-Edge Based Personalized Federated Learning for In-Home Health Monitoring

Qiong Wu, Xu Chen, *Senior Member, IEEE*, Zhi Zhou, *Member, IEEE*, and Junshan Zhang, *Fellow, IEEE*

Abstract—In-home health monitoring has attracted great attention for the ageing population worldwide. With the abundant user health data accessed by Internet of Things (IoT) devices and recent development in machine learning, smart healthcare has seen many successful stories. However, existing approaches for in-home health monitoring do not pay sufficient attention to user data privacy and thus are far from being ready for large-scale practical deployment. In this paper, we propose FedHome, a novel cloud-edge based federated learning framework for in-home health monitoring, which learns a shared global model in the cloud from multiple homes at the network edges and achieves data privacy protection by keeping user data locally. To cope with the imbalanced and non-IID distribution inherent in user's monitoring data, we design a generative convolutional autoencoder (GCAE), which aims to achieve accurate and personalized health monitoring by refining the model with a generated class-balanced dataset from user's personal data. Besides, GCAE is lightweight to transfer between the cloud and edges, which is useful to reduce the communication cost of federated learning in FedHome. Extensive experiments based on realistic human activity recognition data traces corroborate that FedHome significantly outperforms existing widely-adopted methods.

Index Terms—Federated learning, in-home health monitoring, personalization

1 Introduction

With lower fertility and longer life expectancy, population ageing is becoming a global issue. According to the World Health Organization (WHO), the world's population aged 60 years and above will increase to 1.2 billion in 2025 and subsequently to 2 billion by 2050 [1]. The ageing population tends to have a higher prevalence of chronic diseases, physical disabilities, mental illnesses and other co-morbidities, deriving problems such as shortage of medical resources and reduction of quality in healthcare services. Moreover, some older adults (e.g., solitary elderly, elderly couples) prefer to live independently in their own homes, bringing an increased risk of falls and strokes which could be life-threatening. Therefore, there is a growing demand for developing technologies that can aid in the care of the elderly from hospital-centric to home-centric. With the proliferation of smart devices, mobile networks and computing infrastructures, Internet of Things (IoT) is poised to make substantial advances in healthcare systems due to the integrated sensing, computation and communication capabilities of IoT devices. Thus, IoT-based in-home health monitoring is envisioned as a promising paradigm and has attracted great attention [2], [3].

IoT devices, particularly wearable devices as exemplified by smartphones and smartwatches, can track a user's activity, heart rate and so forth persistently, and then feedback these data to a healthcare center for further processing and diagnosis. Due to the easiness and low cost of people's health information acquisition, smart in-home healthcare is able to facilitate pervasive and nonintrusive health monitoring, and is emerging as mainstream for smart healthcare [4]. Current healthcare applications often utilize machine learning (ML) models trained on massive user data to yield insights, perform in-home health monitoring and ultimately provide better public health services and products. As illustrated in Fig. 1, one possible way is to collect a large quantity of training data required for a powerful in-home health monitoring system from many families. As users' health data involves individual privacy, collecting and storing abundant health monitoring data at a remote cloud center to support data analytics based smart healthcare services would impose great privacy leakage risk, making people reluctant to adopt such services. At the same time, with the increasing concerns on data security and user privacy, several regulatory policies and protection mechanisms have been set to restrict data access and protect data privacy. For example, General Data Protection Regulation (GDPR), issued by the European Union, has enforced strict rules on data security and privacy protection [5]. Additionally, China and the United States have also strengthened their attention to data privacy [6]. Under the increasingly stringent data privacy protection legislation, it is almost impossible to integrate sufficient user

- *Qiong Wu, Xu Chen, and Zhi Zhou are with the School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China. E-mail: wuqiong23@mail2.sysu.edu.cn, {chenxu35, zhouzhi9}@mail.sysu.edu.cn.*
- *Junshan Zhang is with the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287 USA. E-mail: junshan.zhang@asu.edu.*

Manuscript received 2 May 2020; revised 18 Nov. 2020; accepted 10 Dec. 2020. Date of publication 16 Dec. 2020; date of current version 1 July 2022. (Corresponding author: Xu Chen.)
Digital Object Identifier no. 10.1109/TMC.2020.3045266

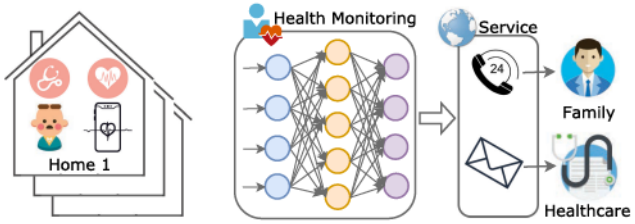


Fig. 1. In-home health monitoring. A large quantity of user data from smart healthcare IoT devices in multiple homes is collected to train machine learning (ML) models for in-home health monitoring and supportive healthcare services. As the user's health data is privacy-sensitive, ML models can not be traditionally managed in a centralized manner.

health data scattered around multiple families for ML model training and in-home health monitoring.

Emerging wisdom to address the above privacy challenge in the traditional, centralized machine learning approaches is federated learning (FL) proposed by Google in 2016. FL enables jointly training a shared global machine learning model under the coordination of a central server by aggregating locally-computed updates while keeping all the sensitive data in local clients (e.g., smartphones), thus allowing users to collectively reap the benefits of the shared model without compromising their data privacy. This form of learning is ideal for community healthcare because it ensures privacy by default, respects data ownership, and maintains locality of data for application deployment at scale. Furthermore, federated learning makes all the participating clients (e.g., medical institutions, smart healthcare IoT devices) share their experiences with privacy guarantee, resulting in a significantly improved performance of the ML model [7].

Nevertheless, existing studies of FL focus on training a single global model, which would suffer from the weaknesses in both statistical and communication perspectives and lacks personalization, resulting in a degraded performance in the healthcare scenarios. First, since different users have different physical characteristics, the personally-generated data naturally exhibits the kind of non-IID (a.k.a. non-independent and identically distributed) distribution. Even for a single user, the health monitoring data can be highly skewed. For example, an adult's activity data may include a lot of standing and walking samples, but very few falls. Both the imbalanced and non-IID distribution of user health data may greatly degrade the learning performance. Second, the communication challenge refers to communication constraints, such as slow or expensive connections and limited bandwidth costs. Existing federated learning approaches mainly focus on the communication rounds reduction considering the fact that mobile devices are frequently offline [8]. However, the communication overload reduction in each round is also critical to improve the efficiency of the algorithm considering the bandwidth cost. Finally, as for personalization, the shared model trained by federated learning only captures the common features of all users, but it may perform poorly on a particular user. Thus, it is necessary to learn the fine-grained information on a particular user for personalized healthcare. However, current federated learning researches in healthcare only focus on one of these challenges, without considering them as an interrelated integration. For example, Zhao *et al.* propose a data-sharing strategy to improve the FedAvg algorithm with non-IID by creating a small globally-shared dataset [9]. Heong *et al.* advocate federated distillation to reduce the communication size to transfer in FL [10]. Chen *et al.* propose a federated transfer learning framework for personalized healthcare [11].

To cope with above issues, we propose FedHome, a cloud-edge federated learning framework for personalized in-home health monitoring, which is a novel scheme that can simultaneously address the three challenges in a holistic manner. FedHome utilizes a synergistic cloud-edge computing architecture and federated learning to avoid uploading users' sensitive data to a centralized cloud server, thus reducing the network latency, and more importantly, addressing privacy concerns. We devise a novel generative convolutional autoencoder (GCAE) network as the model trained on both the cloud and the edges. By integrating the parameter sharing mechanism in convolutional neural network (CNN) and representation ability of autoencoder (AE), GCAE is able to capture the representative and low-dimensional features of user data. Moreover, the representation ability of GCAE is improved together with the prediction ability through the model training process in federated settings. After global model training, GCAE generates new data samples of minority classes in low-dimension space with synthetic minority over-sampling technique (SMOTE) in order to deal with the imbalanced and non-IID distribution exhibited in users' health data. With the generated class-balanced dataset, the trained GCAE model can be personalized by fine-tuning its parameters for precise healthcare performance during classification or inference. It is worthnoting that GCAE plays a significant role in both training and personalization stages, thus resulting in an increased classification accuracy. Furthermore, GCAE is able to reduce communication overhead without compromising performance due to parameter sharing and autoencoder mechanism.

In summary, this paper makes the following contributions:

- We propose FedHome, a novel cloud-edge based personalized federated learning framework for in-home health monitoring. FedHome trains a globally shared model from dispersed homes under the coordination of a central server in a cloud-edge paradigm, which can prevent data leakage by keeping user data locally and then achieve personalized model learning with user's local data.
- FedHome features on a novel design of the generative convolutional autoencoder. By synthesizing samples of minority classes and retraining user's local model with the generated class-balanced dataset, GCAE mitigates the prediction performance degradation caused by imbalanced and non-IID distribution of user data, and achieves superior personalized predictions. Besides, GCAE is a lightweight model by parameter sharing and dimension reduction, which makes the model transfer between the cloud and the edges much more cost-efficient.
- Extensive experiments are conducted using a realistic human activity dataset to demonstrate the effectiveness of FedHome, which shows that FedHome well outperforms traditional centralized methods in both the balanced and the imbalanced data cases. For example, FedHome achieves a high accuracy of

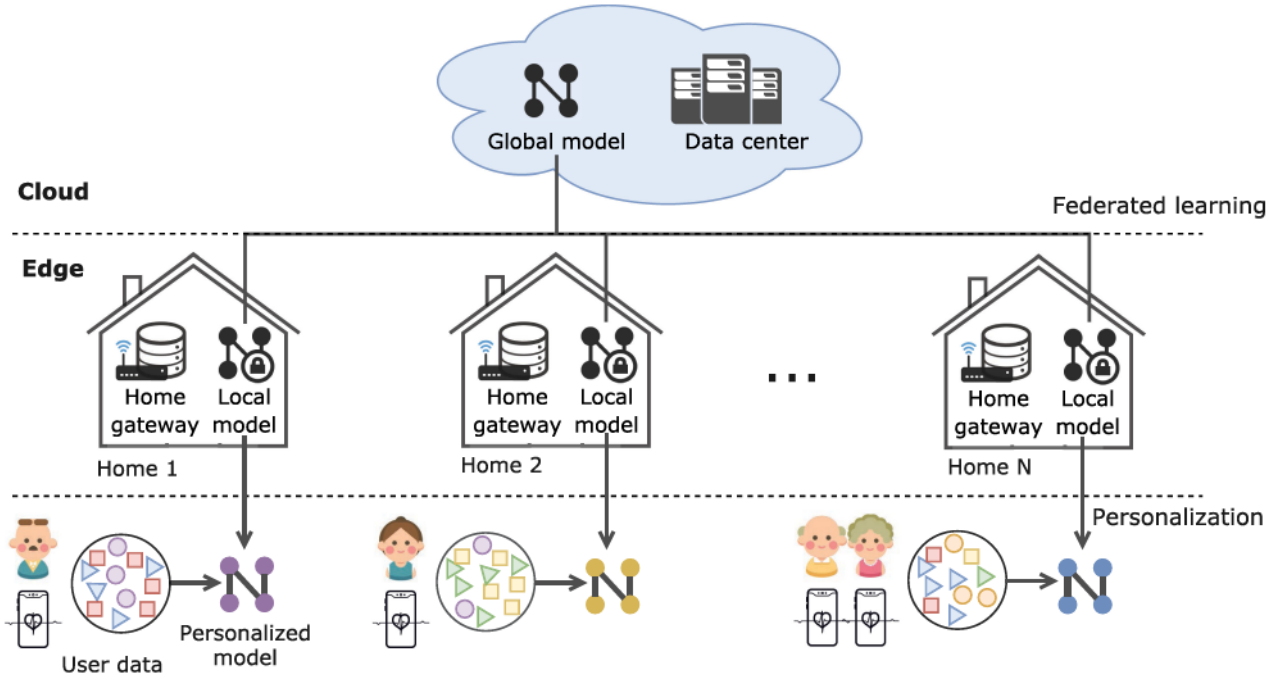


Fig. 2. FedHome framework overview.

95.41 percent with more than 7.49 percent accuracy improvement over the widely-used convolutional neural network algorithm. Comparing with other federated learning approaches, FedHome achieves a gain of more than 10 percent accuracy.

The rest of this paper is organized as follows: Section 2 outlines the framework overview of FedHome. In Section 3, we introduce the algorithm details of FedHome. In Section 4, we conduct extensive experiments using a realistic human activity dataset. In Section 5, we review the related work in in-home health monitoring and federated learning. Finally, we conclude our paper in Section 6.

2 Overview of FedHome

We consider an in-home health monitoring system consisting of a cloud server and N dispersed edge computing nodes (e.g., smarthome gateway) deployed at users' homes. Given a specific task, for example, elderly fall-down detection via human activity recognition, we wish to train an accurate machine learning model by taking advantage of the abundant data $\{\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_N\}$ collected from all the home users in N edge nodes. Conventional approaches usually put all the data together (e.g., uploading to a cloud server) to train a model $\mathcal{M}_{ALL}$ in a centralized fashion. However, such approaches would cause significant privacy issues as the health data usually contains a user's sensitive private and personal information. Thus, we resort to the federated learning paradigm, in which users collaboratively train a model $\mathcal{M}_{FED}$ without uploading their personal data to a central cloud server or exposing to each other.

There have been some attempts that study healthcare in federated learning settings. For example, Silva *et al.* present a federated learning framework by investigating brain structural relationships across diseases and clinical cohorts without sharing individual information [12]. Kim *et al.* convert massive electronic health records into meaningful phenotypes for data analysis by iteratively transferring secure summarized information of hospitals to a central server in federated settings [13]. Liu *et al.* develop a two-stage federated natural language processing method that enables utilization of clinical notes from different hospitals or clinics without moving the data and demonstrate its superior performance [14]. Besides, healthcare giants around the world—including the American College of Radiology (ACR), MGH and BWH Center for Clinical Data Science, and UCLA Health—are piloting FL in healthcare by bringing AI with privacy to hospitals [15]. For example, ACR uses FL to allow its medical imaging members to securely build, share and adapt ML models. Partners HealthCare, an integrated health system founded by Brigham and Women's Hospital and Massachusetts General Hospital, has also announced a new initiative using the NVIDIA Clara FL framework that will combine technical and clinical expertise with privacy preservation to deliver more robust AI algorithms for healthcare from hospital to home [16]. Nevertheless, these studies of FL for healthcare applications pose new challenges from both statistical and communication perspectives and lacks personalization.

Specifically, in this paper we propose FedHome, a novel framework that aims to achieve accurate in-home health monitoring without compromising user data privacy by leveraging the merits of federated learning and edge computing. As shown in Fig. 2, FedHome adopts a cloud-edge architecture for distributed data processing at users' homes, achieves collaborative model training through federated learning and performs personalized model application with user's local data. We elaborate these key elements in detail as follows.

Cloud-Edge Architecture. As elaborated above, purely cloud-based health monitoring would risk at the privacy leakage since it requires users to upload their sensitive health related monitoring data. On the other hand, in some

emergency situations (e.g., fall detection for elderly people), the cloud-based approach could fail to respond in real time due to the significant network latency or interruption over the Internet. To address these issues, we promote a synergistic cloud-edge architecture to bring necessary on-demand edge computing power in the proximity of IoT devices. Therefore, each user can choose to process its local data on device, or offload its intensive model training tasks to the trustworthy edge nodes (e.g., edge gateway at home) for fast training. By leveraging federated learning, all users can merge their local models into a global model in the cloud for knowledge sharing without exposing their sensitive data. The trained model deployed locally at the edge can conduct inference much faster than in the cloud, and hence is useful for providing real-time healthcare services. Note that the cloud server here can be provided by some healthcare service company or can be a public server deployed collectively by the entire community of the home users.

Furthermore, as family members in the same home trust each other, they can share data and train a family-shared model at the trustworthy edge nodes, thus family members with insufficient private data can develop accurate local models by reaping the benefits from the data of their mutual-trusting family members.

Federated Learning. The federated learning procedure of FedHome mainly consists of the following three stages. Indicated by recent work that local models trained from different initial conditions would perform poorly in federated learning [17], we initialize the cloud model and then send the initial state to all clients (homes participated in FL) so that client models can be trained from the same random initialization. After receiving the global cloud model, each client model at the edge then performs local computation based on the global model and its local data, and sends the updated local model parameters to the cloud. Finally, the cloud server aggregates model updates submitted by participant clients and averages these updates into its global model. The process repeats until it converges after many rounds of iterations. Note that all these steps do no share any user data or information but the model parameters.

Personalization. The trained cloud model is based on generic datasets from dispersed homes, which may not well capture the specific characteristics of an individual target user. To perform personalized in-home health monitoring, each user can train a personalized model by integrating the trained global model and her personal health data. Nevertheless, as user's personalized data is usually insufficient and high-skewed, there may be a significant distribution difference between user's local data and the data population of all the participating clients, leading to large model weight divergence between user's personalized model and the global model [9]. Thus, directly retraining the cloud model based on user's local data may result in an even worse model (e.g., degraded accuracy, overfitting with a few data samples) [18]. To tackle this problem, the model trained in the cloud and the edges should be carefully designed. Specifically, we devise a generative convolutional autoencoder network to deal with the imbalanced class distribution of user's health monitoring data. Besides, GCAE is suitable to be trained under federated learning as it is lightweighted and can reduce the communication cost between the cloud and the edges. The details of GCAE are elaborated in Section 3.2.

TABLE 1
List of Key Notations

Symbol	Description
N	Total number of clients in a healthcare system
K	Number of activated clients participating in a communication round
$\mathcal{D}_k$	Available dataset on client k
D_k	The set of indexes of dataset on client k
n	Total number of data samples participating in a communication round
n_k	Number of data samples available from client k
E	Number of training passes over its local dataset for a client in a communication round
B	Local minibatch size used for client updates
η	learning rate of federated learning
λ	weighting factor for the decoder network

3 Learning Algorithm for FedHome

In this section, we describe the learning algorithm for FedHome, which consists of personalized federated learning and generative convolutional autoencoder. The key notations used in this paper are summarized in Table 1.

3.1 Personalized Federated Learning

Federated learning, which enables many participating clients to train a shared global model without sharing their private data, is the core building block for the FedHome framework. In FedHome, both the cloud model and the client models at the edges are learned by deep neural networks (DNNs) as DNNs have been identified as a universal function approximator with excellent generalization and approximation capabilities [19]. Let F_k and D_k denote the local empirical objectives and the set of indexes of local data on client k (edge server k), respectively. n_k is the number of samples available from client k and $n = \sum_{k=1}^{K} n_k$ is the total sample size in a communication round, where K is the number of active clients participating in federated learning. Thus, federated learning problem boils down to solving an empirical risk minimization problem of the form [17]

$$\min_{\omega \in \mathbb{R}_d} f(\omega) = \sum_{k=1}^{K} \frac{n_k}{n} F_k(\omega) \quad \text{where} \quad F_k(\omega) = \frac{1}{n_k} \sum_{i \in D_k} f_i(\omega). \tag{1}$$

The objective $f(\omega)$ in federated learning can be rephrased as a linear combination of the local empirical objectives $F_k(\omega)$. The learning objective is task-specified, for instance, the objective can be cross-entropy loss for classification tasks. ω denotes the parameters to be learned, i.e., the weights and bias in our deep neural networks.

`FederatedAverage` (FedAvg) algorithm, which combines local stochastic gradient descent (SGD) on each client with a server that performs iterative model averaging, serves as the fundamental framework in federated learning settings [17]. In each communication round t, each edge client k computes $g_k = \nabla F_k(\omega_t)$, the average gradient on its local data at the current model parameters ω_t, and then the

cloud server aggregates all these gradients submitted from clients and updates the global model by

$$\omega_{t+1} \leftarrow \omega_t - \eta \sum_{k=1}^{K} \frac{n_k}{n} g_k, \tag{2}$$

since $\sum_{k=1}^{K} \frac{n_k}{n} g_k = \nabla f(\omega_t)$. As communication cost is expensive in federated learning, FedAvg proposes to use additional computation in order to decrease the number of communication rounds. Two primary ways to achieve this are outlined as follows: 1) increasing parallel clients participating in each communication round; and 2) increasing computation on each client by iterating the local update $\omega^k \leftarrow \omega^k - \eta \nabla F_k(\omega^k)$. In each communication round, the amount of computation for each client is controlled by E and B, which represent for the number of training passes over its local dataset and the local minibatch size, respectively.

FedAvg has been demonstrated to be accurate and robust in image classification tasks and language modeling tasks [17]. In this paper, we adopt FedAvg algorithm for cloud-edge collaborative training in FedHome. Note that, to further enhance the data privacy for the in-home health monitoring scenario, the above parameter exchange and gradient aggregation steps are done with homomorphic encryption [20], [21], which is a powerful tool to secure the learning process by computing on encrypted data and has been studied and adopted in FL settings. For example, Federated AI Technology Enabler (FATE) initiated by WeBank's AI division is an open-source technical framework to support federated AI ecosystem that enables secure computing protocols based on homomorphic encryption (HE) and multi-party computation [22]. The learning process of FedHealth framework can leverage homomorphic encryption to avoid information leakage of model parameters during the learning process [11]. To reduce the computational cost, more efficient variants of HE, such as additive HE, can be adopted in practical FL systems rather than fully homomorphic encryption which is computationally expensive [23].

After the global cloud model is learned, it can be directly applied to the homes at the edges for in-home health monitoring. However, it is obvious that the distribution of training samples for the cloud model is highly different from that generated by a single user, thus the cloud model fails in personalization. To make the model more tailored to a specific user, each edge client at home can train a personalized model by integrating the cloud model and the health monitoring data generated by its own users. Nevertheless, as user's personalized data is usually insufficient and high-skewed, there may be a significant distribution difference between user's local data and the data population of all the participating clients, leading to large model weight divergence between user's personalized model and the global model [9]. Thus, directly retraining the cloud model with user's local data may result in an even worse model (e.g., slow convergence rate, degraded accuracy and overfitting with a few data samples) [18]. To cope with this issue, we devise a generative model called GCAE under federated learning paradigm for better personalization, which is elaborated in the following.

3.2 Generative Convolutional Autoencoder

The health monitoring dataset stored in the edge server of home k can be defined as $\mathcal{D}_k = \{(\mathbf{x}_i, y_i)\}_{i=1}^{n_k}$. The feature $\mathbf{x}_i$ contains the health information and is always high-dimensional due to a multitude of monitoring signals for healthcare, and the distribution of class label y_i is usually skewed and imbalanced since abnormal states or events such as elderly falling-down are usually sparse during the entire monitoring process. To deal with the imbalanced data distribution issue, we propose to learn a generative prediction model in both the cloud and the edges. More specifically, we devise a generative convolutional autoencoder that utilizes autoencoder as the backbone. AE aims to learn compact and effective feature representations of input values via an unsupervised learning schema by mapping the input to a latent space through encoder network and then mapping back to the original space using decoder network [24]. The learned feature representations are usually low-dimensional and contain all the information needed to recover the original inputs, and thus can be used as feature vector inputs to a supervised learning model (e.g., Multi-Layer Perceptron–MLP) for user health monitoring. It's worthnoting that the feature representations, which are learned by compressing useful information of the local input data into low dimension by the encoder network of GCAE, are always learned and maintained by the local user without submitting to the cloud. And the cloud server only aggregates the model parameter updates obtained by performing local computations based on the global model and the local data of home users, without access to user's sensitive data (e.g., data samples, low-dimensional feature representations), hence user's privacy can be well protected in our FedHome framework. Besides, due to the fact that health monitoring data is mostly medical images and sensor records, we use convolutional neural network as the main architecture for both encoder and the decoder. CNN is competent to extract features from signals and it has achieved promising results in image classification and speech recognition [25].

Fig. 3 depicts the details of GCAE, which consists of a convolutional pyramid encoder e_ϕ that focuses on learning low-dimensional, common and representative features of high-dimensional inputs $\mathbf{x}$, a convolutional decoder d_θ that calculates the reconstruction loss between $\mathbf{x}_i$ and $d_\theta(e_\phi(\mathbf{x}_i))$

$$\mathcal{L}_d((\phi, \theta), \mathcal{D}_k) = \sum_{i=1}^{n_k} ||\mathbf{x}_i - d_\theta(e_\phi(\mathbf{x}_i))||_2^2, \tag{3}$$

and a Multi-Layer Perceptron (MLP) p_ψ that predicts the likelihood of the class label y_i. The objective of MLP is to minimize the difference score between the actual label and predicted probability distributions for all classes

$$\mathcal{L}_p((\phi, \psi), \mathcal{D}_k) = \sum_{i=1}^{n_k} loss(y_i, p_\psi(e_\phi(\mathbf{x}_i))), \tag{4}$$

where the loss function $loss(\cdot, \cdot)$ denotes the cross-entropy loss for specific prediction tasks, e.g., human activity recognition task. As we wish our health monitoring model is equipped with the generation ability, we propose to learn the representative low-dimensional representations and conduct prediction jointly. Thus, we train an end-to-end model by minimizing

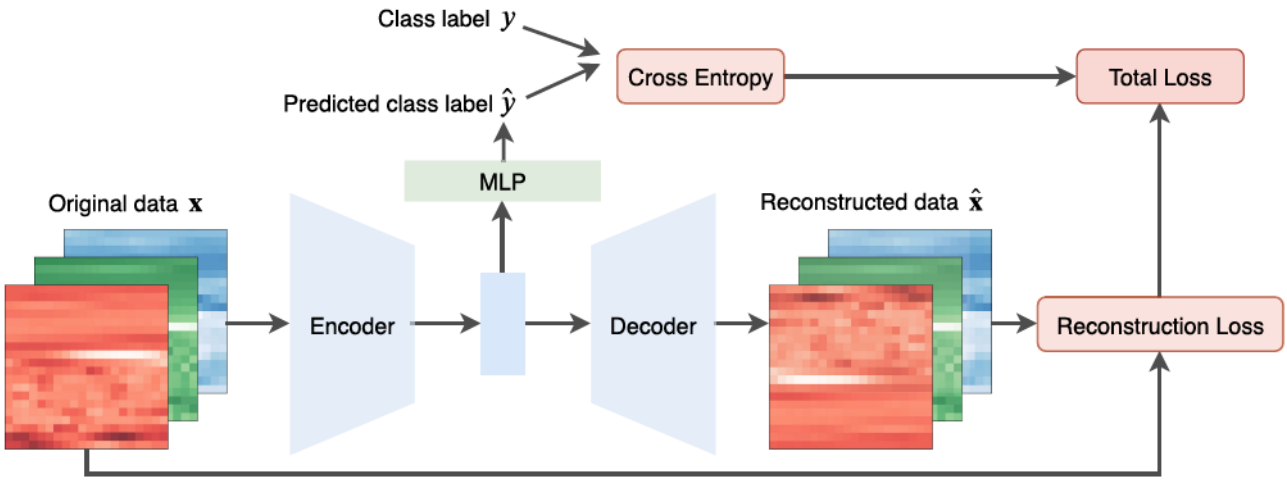


Fig. 3. Architecture of the generative convolutional autoencoder.

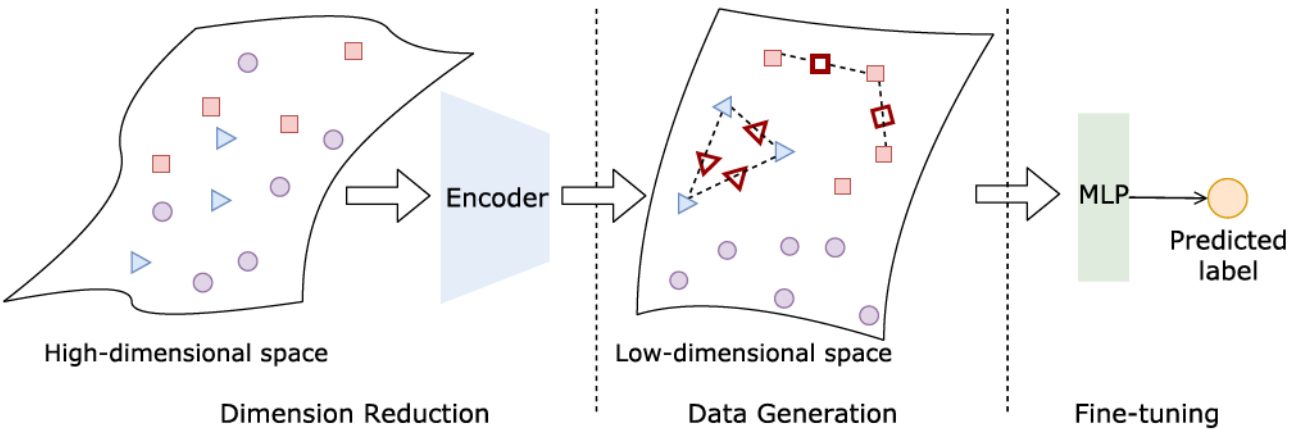


Fig. 4. Before inference or classification, the health data is mapped to a low-dimensional space to generate samples in minority classes with SMOTE. With the synthetic class-balanced dataset, we can fine-tune the prediction network for more accurate health monitoring.

$$\mathcal{L}(\omega, \mathcal{D}_k) = \mathcal{L}_p + \lambda \mathcal{L}_d, \tag{5}$$

where $\omega = (\phi, \theta, \psi)$ stands for the parameters of the three models and $\lambda \in \mathbb{R}^+$ is the weighting factor that controls the impact of the decoder network on the whole model.

In the training process of FedHome, the cloud server managed by smart healthcare service provider first initializes the GCAE model and then sends the initialized model to all the participating clients (e.g., smartphones). Then, each IoT device user can offload its GCAE training task to the trustworthy edge (e.g., smart gateway or edge server) at home for fast computation. Finally, a global GCAE model can be well trained by leveraging model updates from multiple home edges under the coordination of the central cloud server in a cloud-edge architecture.

Data Generation. When the learned GCAE model is deployed to the edge, we first conduct personalization to make the model more tailored to the edge users. However, the dataset generated by a single user may suffer from data imbalanced problem, which may heavily impact the model performance as a well-trained deep neural network usually relies on a class-balanced training dataset. One common way to tackle the issue of imbalanced data is over-sampling which aims to increase the number of instances from the underrepresented classes in the dataset.

Due to the high dimensional and complex characteristics of health monitoring data, it is hard or even impossible to conduct over-sampling. As encoder network e_ϕ of GCAE is able to capture representational, common and low-dimensional features of original input $\mathbf{x}_i$, we propose to conduct over-sampling in the set of $e_\phi(\mathbf{x}_i)$ instead of $\mathbf{x}_i$. In our design, the encoder network of GCAE encodes the high-dimensional input data $\mathbf{x}_i$ with dimension of $200 \times (3+3) = 1200$ to low-dimension (e.g., 200 in our experiments). Thus, the dimension reduction ratio is 6 in our studying case. It is flexible to determine the dimension reduction ratio according to the specific healthcare applications. Then we adopt synthetic minority over-sampling technique, one of the most commonly used oversampling methods, to generate a class-balanced dataset [26]. As depicted in Fig. 4, SMOTE first randomly selects one or more of the k-nearest neighbors for each instance in the minority classes and then adopts linear interpolation to generate new instances. A salient advantage of over-sampling in low dimension space is that the accuracy in mimicking the distribution of data samples can significantly improve when the dimension reduces [27]. Thus, a class-balanced dataset in low-dimensional space can be obtained based on a given user's original dataset.

Personalization. As the cloud model only learns the coarse features from all users, we focus on learning the fine-grained information of a particular user based on the reconstructed

class-balanced dataset for personalization. Specifically, usually the features in deep networks are highly transferable in lower layers of the network since they focus on learning the common and low-level features, and hence encoder network can be directly reused in the personalized model [28]. While the higher layers learn more specific features to the task and the user, which means that we should refine the prediction model p_ψ (i.e., Multi-Layer Perceptron–MLP) since it is at the higher layers of GCAE. We fine-tune the model parameters of p_ψ by retraining with the reconstructed class-balanced dataset based on user's personal data. In this way, each user can obtain a more accurate personalized model for in-home health monitoring.

Algorithm 1. The Learning Procedure of FedHome

Input: Dataset from N distributed clients $\{\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_N\}$, participating client number K in each communication round, local minibatch size B, number of local epochs E, and learning rate η.

/*federated learning process*/
Cloud server executes:
  Construct a GCAE model $\mathcal{M}_S(\omega) = \{e(\phi), d(\theta), p(\psi)\}$ as the cloud model and initialize its model parameters $\omega_0 = \{\phi_0, \theta_0, \psi_0\}$
  for each round $t = 0, 1, 2, \ldots$ **do**
    randomly select active client set U_t with size K from N clients
    distribute $\mathcal{M}_S(\omega_t)$ to all clients in U_t via homomorphic encryption
    for each client $k \in U_t$ **in parallel do**
      $\omega_{t+1}^k \leftarrow$ **ClientUpdate**(k, ω_t)
    end for
    upload all client models to the server using homomorphic encryption
    /*update cloud model by averaging client models*/
    $\omega_{t+1} \leftarrow \sum_{k \in U_t} \frac{n_k}{n} \omega_{t+1}^k$
  end for
  get the learned global model parameter ω_S
  distribute the learned model $\mathcal{M}_S(\omega_S)$ to all clients via homomorphic encryption
  for each client $k = 1, 2, \ldots, N$ **do**
    $\mathcal{M}_k(\omega^k) \leftarrow$ **ClientPersonalization**$(\mathcal{D}_k, \omega_S)$
  end for

Output: cloud model $\mathcal{M}_S(\omega_S)$ and personalized user model $\mathcal{M}_k(\omega^k)$, $k \in \{1, 2, \ldots, N\}$.

Communication Overhead Reduction. GCAE model not only has the ability to generate new data, but also reduces communication overhead to some extent. To well capture the health information and insights behind the data, it is often necessary to train a large neural network which has millions of parameters. However, in federated setting, a lightweight model is highly desirable in order to reduce the communication overhead during model parameter updates among the cloud and the edges.

Parameter sharing scheme is widely-used to reduce the number of training parameters. In our design, both the encoder and decoder networks of GCAE are convolutional neural networks in which parameter sharing is used. More concretely, convolutional layer, the core building block of CNN, shares weights by all neurons in a particular feature map (the output received on convolving the image with a particular filter). Max-pooling layer, followed convolutional layer, is to progressively reduce the spatial size of the representation. However, purely adopting CNN will cause a lot of information loss and further lead to inaccurate health monitoring. With the autoencoder mechanism, the encoder network of GCAE is able to learn compact and representative feature vector of user's health monitoring data via unsupervised scheme, ensuring that the low-dimensional input feature vector $e_\phi(\mathbf{x}_i)$ fed in prediction network is nearly lossless. Therefore, GCAE is able to reduce communication overhead without compromising performance, which is of great significance in federated settings.

Algorithm 2. Model Update on Each Edge Server

ClientUpdate(k, ω): // Run on client k
  $\mathcal{B} \leftarrow$ (split $\mathcal{D}_k$ into batches of size B)
  for each local epoch e from 1 to E **do**
    for batch $b \in \mathcal{B}$ **do**
      $\omega \leftarrow \omega - \eta \nabla \mathcal{L}(w, b)$
    end for
  end for
  return the trained model parameter ω

Algorithm 3. Personalization on Each Edge Server

ClientPersonalization$(\mathcal{D}, \omega)$: //Personalization process
  calculate low-dimensional feature representation $e_\phi(\mathbf{x}_i)$ for $\mathbf{x}_i$ using encoder network of GCAE for all samples in $\mathcal{D}$ and get the low-dimensional dataset $\mathcal{D}'$
  construct a class-balanced dataset $\mathcal{D}'_b$ by adopting SMOTE algorithm on $\mathcal{D}'$
  refine model parameters of prediction network $p(\psi)$ with dataset $\mathcal{D}'_b$ and get a new prediction model $p(\psi')$
  return model $\mathcal{M}(\omega') = f(\phi, \theta, \psi')$

3.3 The Holistic Algorithm

To sum up, the holistic mechanism of FedHome is presented in Algorithm 1, in which the edges train a globally shared GCAE model for in-home health monitoring under the coordination of a cloud server by leveraging federated learning. In each communication round, each edge uses the current cloud model as the initial model parameters in order to train its local model using its local user data, and then sends its local model parameters to the cloud (see Algorithm 2). This process repeats until the GCAE model on the cloud converges. Then, the trained GCAE model can be deployed in the edge server of each home. Each home user can further learn a personalized model by synthesizing a class-balanced dataset with her personal data and then refining model parameters with the generated dataset (see Algorithm 3). Note that this framework works continuously with the new emerging user data. In other words, FedHome is able to perform incremental learning [29]. That is, when facing new user data, both cloud model and edge client models can be updated continuously. Moreover, the longer the user uses the healthcare application, the more personalized the client model can be. Besides, the learned cloud model, which captures the generic information of healthcare application users, can be easily deployed as a prior

TABLE 2
Human Activity in MobiAct Dataset

Category	Code	Activity	Description
ADLs	STD	Standing	Standing with subtle movements
	WAL	Walking	Normal Walking
	STU	Stairs up	Stairs up (10 stairs)
	STN	Stairs down	Stairs down (10 stairs)
	JUM	Jumping	Continuous jumping
	JOG	Jogging	Jogging
Fall-like Activities	CSI	Car step in	Step in a car
	CSO	Car step out	Step out of a car
	SCH	Sit chair	Sitting on a chair
Falls	Fall	Forward-lying	Fall forward from standing, use of hands to dampen fall
		Front-knees-lying	Fall forward from standing, first impact on knees
		Sideward-lying	Fall sideward from standing, bending legs
		Back-sitting-chair	Fall backward while trying to sit on a chair

model to a new joining home that is new for the healthcare application and has almost no usage records.

The FedHome framework can also adopt other deep neural networks tailored to specific task, for example, recurrent neural networks for modeling Alzheimer's Disease (AD) progression [30], attention network for dementia status prediction from brain magnetic resonance imaging [31]. This makes FedHome as a general framework flexible for supporting many privacy-preserving healthcare applications.

4 Experiments

In this section, we first briefly describe the used dataset and our implementation details for human activity recognition, a specific task highly relevant to in-home health monitoring. Then we evaluate the performance of FedHome framework by comparing with both traditional widely-adopted approaches and federated learning based methods.

4.1 Dataset Description and Preprocessing

Our study in this paper is based on a publicly accessible human activity recognition dataset called MobiAct [32]. There are 57 volunteers (42 men and 15 women) within an age bracket of 20-47 years participating in the generation of MobiAct dataset. Each volunteer wears a Samsung Galaxy S3 smartphone with the accelerometer and gyroscope sensors. The tri-axial linear accelerometer and angular velocity signals are recorded at a constant sampling frequency of 200 Hz by the embedded sensors while volunteers perform predefined activities. The recorded activities can be divided into three types: (1) activities of daily living (ADL): the most common everyday activities like walking, standing, stairs up and stairs down; (2) fall-like activities which are sudden or rapid and are similar to falls, such as sitting on a chair or stepping in and out of a car; (3) falls, for example, fall forward from standing, fall backward while trying to sit on a chair, etc. Typically, there are four different types of falls as described in Table 2. Note that this dataset can provide the relevant application scenario to mimic in-home health monitoring through human activity recognition (e.g., fall detection for elderly people).

Before data preprocessing, we first introduce human activity model which is established based on the Cartesian coordinate system [33] in accordance with the direction of the sensors in a smartphone as shown in Fig. 5a. In the 3-dimensional coordinate system, a_x, a_y, a_z denote the acceleration along the x, y and z axis, respectively. Correspondingly, ω_x, ω_y, ω_z represent for the angular velocity of the human body around the x, y and z axis. Since one second is enough for users to perform an activity (e.g., from falling down to touching the ground), a fixed 1-second sliding window is used for feature extraction and the overlapping rate between successive windows is 80 percent.

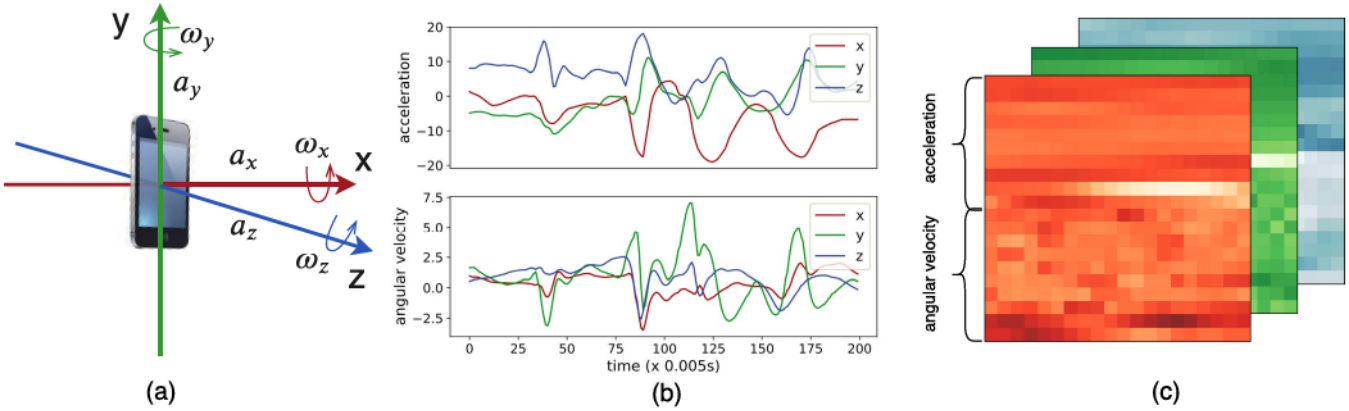


Fig. 5. Data preprocessing procedure. (a) Human activity model. (b) Tri-axial acceleration and angular velocity data segmented for one activity. (c) Mapping the tri-axial sensor data into an RGB image.

Fig. 5b shows the tri-axial acceleration and angular velocity data segmented for one activity. As the sampling frequency is 200 Hz, each activity contains 200 data points for each sensor along each axis, and totally $200 \times (3+3) = 1200$ data points. If we consider the 3-axes of the human activity model as the 3-channels of an RGB image, the values of the tri-axial acceleration and angular velocity data can be mapped into the pixels in an RGB image. As depicted in Fig. 5c, the first 200 data in the RGB image are tri-axial accelerations while the last 200 data are tri-axial angular velocities.

To practically mimic smart healthcare environments, we randomly select 30 volunteers and regard them as users in multiple dispersed homes. Each user can offload their model training task to the trustworthy edge node (e.g., smart gateway or personal computer at home) for fast training or train the model locally on device. Moreover, as the skewed data distribution is an inherent characteristic in federated settings, we employ three approaches of training data partition:

- *Balanced data partition:* There are 30 homes with one isolated user at each home. For each user, we extract 48 samples for each activity. Totally, one user has 480 samples for model training.
- *Imbalanced data partition:* There are 30 homes with one isolated user at each home. For each user, we take a random number of samples for each activity and finally, we ensure that each user has 480 samples for model training.
- *Home data partition:* The dataset for each user is the same with imbalanced data partition case. However, we randomly divide these 30 users into 10 homes with 1 to 5 family members per home. Users can share their health data with their family members at the edge.

The test data for the above three data partition manners is the same. For the selected 30 users, each user has 160 samples under balanced distribution. It is worth noting that the data distribution is always non-IID as different users have different physical characteristics, lifestyles and lifelong medical data, and as a consequence, different data distributions.

4.2 Implementation Details

Generative convolutional autoencoder network is adopted on both the cloud server and the edges at users' homes for model training and prediction. The encoder network of GCAE is composed of 3 convolutional layers with filter size of 3×3 to encode the original data into a hidden latent representation. The filter numbers of the three convolutional layers are 32, 16 and 8, respectively, and each of the first two convolutional layers is followed by a 2×2 max-pooling layer. For the decoder network, three 5×5 deconvolution layers with 16, 32, 3 filters and two 2×2 up-sampling layers are applied to decode the hidden representation. The activation function for the third convolutional layer is *Sigmoid* and we use *Relu* function for other layers in both the encoder and the decoder networks. The output of the encoder is flattened and then fed in an MLP network which is composed of 2 fully-connected layers with 128 and 10 units, respectively. *Softmax* is adopted in the final fully-connected layer of the MLP network to calculate the probability of the output results. The weighting factor of the decoder network, λ, is set to be 0.01. The model is trained by minibatch Stochastic Gradient Descent optimizer with a learning rate of 0.01. These hyperparameters are tuned using cross-validation procedure.

4.3 Experimental Results

4.3.1 Baselines

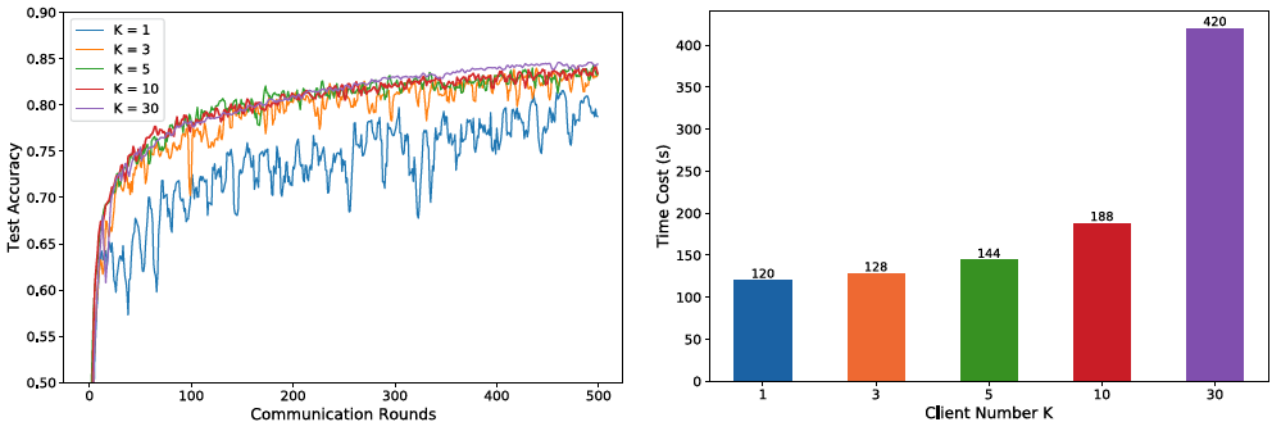
The baselines for human activity recognition can be divided into the following two groups:

Traditional Centralized Algorithms. traditional algorithms usually collect a large amount of user data in a centralized cloud server to train a satisfactory model. Support Vector Machine (SVM), k-Nearest Neighbor (kNN) and Random Forest (RF) are widely adopted in many healthcare applications [34]. With the recent success of deep learning, Multi-Layer Perceptron and Convolutional Neural Network are also introduced to in-home health monitoring and have achieved great performance improvement [35]. We also test the accuracy of Generative Convolutional Autoencoder proposed in this paper in a centralized fashion. We collect all the training data of 30 users to form the training dataset of traditional centralized algorithms and describe these algorithms as below.

- *Support Vector Machine (SVM)* is a popular classifier that is demonstrated to be effective on a wide range of classification problems [36].
- *k-Nearest Neighbor (kNN)* is a non-parametric method. It is introduced to measure the difference or similarity between instances according to a distance function [37].
- *Random Forest (RF)* is an ensemble learning method that builds a set of decision trees with random subsets of attributes and then bags them for classification results [38].
- *Multi-Layer Perceptron (MLP)* is a supervised learning technique that makes final predictions through fully-connected layers. In our design, the MLP network is composed of three fully-connected layers with 1200, 100 and 10 neural units.
- *CNN* has strong representation capability and is widely adopted in healthcare scenarios, such as in human activity recognition [39]. CNN used in this paper is designed as the combination of an encoder and an MLP classifier which has the same architecture with that in our GCAE model.
- *Generative Convolution Autoencoder (GCAE)* takes advantage of the feature extraction ability of CNN, dimension reduction capability of autoencoder and over-sampling strategy of SMOTE [40].

Federated Learning Based Methods. in federated settings, each client trains a local model with its personal-generated data, and FedAvg method is applied to obtain a globally shared model by aggregating local models without compromising users' privacy. This process repeats until it converges.

- *FL-MLP:* both the cloud and client models are implemented with MLP.
- *FL-CNN:* CNN model is adopted as both the cloud and client models.
- *FL-CNN-Large:* The difference between FL-CNN-Large and FL-CNN is that the filter numbers of the three convolutional layers in the CNN model of FL-CNN-Large are 64, 32 and 16, respectively.



(a) Test Accuracy (b) Time Cost

Fig. 6. The test accuracy and time cost under different number of participating clients in each communication round. We choose K = 5 by making a trade off between the stability and the efficiency of the learning algorithm.

- *FedHome:* FedHome is the proposed federated learning framework for in-home health monitoring. GCAE is used as both the cloud and the client models. When the training process converges, the client will first balance its data distribution and then conduct model personalization with its own personal data.
- *FedHome-p:* FedHome framework without client personalization is named FedHome-p.

4.3.2 Performance Evaluation

In federated settings, there are three key parameters which may influence the performance of federated learning: (1) K, the number of participating clients (homes participating in FL) in each communication round; (2) E, training passes over each client's local data and (3) B, local minibatch size. We first conduct experiments with the number of clients participating in each communication round. First, we fix $B = 10$ and $E = 5$, and evaluate the test accuracy under different values of K. More concretely, we study the effect of K by setting K equal to 1, 3, 5, 10 and 30, which can also be regarded as $\frac{1}{30}$, $\frac{1}{10}$, $\frac{1}{6}$, $\frac{1}{3}$ and 100 percent of the whole training data. As shown in Fig. 6a, the learning algorithms with different choices of K can all converge within 500 communication rounds. The test accuracies under different values of K improve with the increase of the communication rounds. However, when K is small (e.g., $K = 1$), the test accuracy shows a significant gap compared with the case with bigger K. For example, the test accuracy is 79.15 percent when $K = 1$ while 84.23 percent in the case that $K = 30$. Besides, the learning curve exists erratic fluctuation to some extent with a smaller K and becomes smoother as K increases. Although the test accuracies are similar except the case that $K = 1$, the training time for each value of K varies dramatically as demonstrated in Fig. 6b. For example, the training time in the $K = 1$ case is 120s when experimenting on a desktop PC equipped with a quad-core Intel processor at 3.4 GHz with 8 GB of RAM. The training time for $K = 3$ is 1.07 times longer compared with the training time when $K = 1$, while the training time for $K = 30$ is 3.28 times more than that in the $K = 3$ case. We make a trade-off between the stability and the efficiency of the learning algorithm in the training process and fix $K = 5$ for the following experiments.

To unfold the impact of parameter B and E on test accuracy and communication cost, we set six combinations of parameter settings as shown in Fig. 7. We can see that the learning curve converges rapidly with larger E and smaller B. When $B = 10$ and $E = 20$, the test accuracy of human activity recognition can reach 83.57 percent within only 300 communication rounds. While for the case that $B = 50$ and $E = 1$, it may require a large number of communication rounds to reach the same test accuracy. In fact, for a client with n_k local samples, the number of local updates per round is given by $E\frac{n_k}{B}$. Thus, either increasing E, decreasing B, or both, will result in more computation per client on each round. This observation inspires us that adding more local SGD updates per communication round can produce a dramatic decrease in communication cost, which is essential in federated learning. As the performance of the case $B = 10$ and $E = 5$ is comparable with the best one, we choose to use $B = 10$ and $E = 5$ under the consideration of computation cost for the following experiments. It's worthnoting that the training process of FedHome can be conducted offline to learn a satisfactory shared model by reaping the benefits of user data from multiple homes without compromising their data privacy. When the globally shared model is deployed on edges or user's IoT devices, FedHome enables fast personalization (retraining with local data of a family or a user) and real-time inference (prediction with

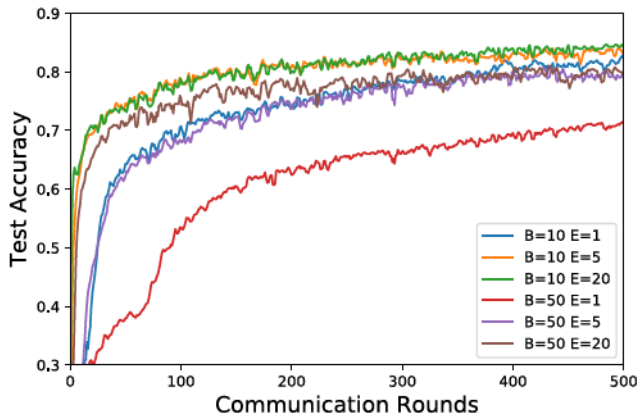


Fig. 7. Test accuracy versus communication rounds under different combinations of B and E.

TABLE 3
Test Accuracy of the Models in Human Activity Recognition

Methods	Test Accuracy		Model Parameters
	Balanced Data	Imbalanced Data	
SVM	$77.25 \pm 1.77\%$	$67.88 \pm 0.50\%$	–
KNN	$80.85 \pm 1.48\%$	$74.99 \pm 0.67\%$	–
RF	$84.27 \pm 0.34\%$	$74.28 \pm 0.29\%$	–
MLP	$92.31 \pm 1.45\%$	$87.94 \pm 1.32\%$	–
CNN	$91.77 \pm 1.43\%$	$87.92 \pm 0.15\%$	–
GCAE	$92.02 \pm 0.92\%$	$88.10 \pm 0.14\%$	–
FL-MLP	$89.28 \pm 1.17\%$	$85.06 \pm 0.59\%$	$1,562,310\,(29.96\times)$
FL-CNN	$85.07 \pm 0.44\%$	$82.91 \pm 0.34\%$	$33,698\,(0.65\times)$
FL-CNN-Large	$87.24 \pm 0.46\%$	$84.87 \pm 0.72\%$	$77,498\,(1.49\times)$
FedHome-p	$89.13 \pm 0.93\%$	$84.22 \pm 0.37\%$	52,149
FedHome	$\mathbf{95.87 \pm 0.23\%}$	$\mathbf{95.41 \pm 0.11\%}$	**52,149**

the input data by simply conducting matrix multiplications of the local neural network model).

Table 3 illustrates the performances of different human activity recognition approaches under both balanced and imbalanced data cases. For each method, we compute the average and standard deviation of test accuracy by repeating the training and prediction processes five times. We can see that the test accuracies of all methods in balanced dataset are higher than those in imbalanced dataset. Moreover, the results show that deep learning based methods (MLP, CNN and GCAE) can all achieve a higher accuracy than traditional machine learning methods (SVM, KNN and RF) when the data samples of all clients are aggregated in a centralized location. This is due to the strong representation capability of deep neural networks. Besides, deep learning based methods can be updated and enhanced by incremental learning without retraining. This property is valuable in federated learning where model reuse is important and helpful because new data emerges continuously and the model should be quickly adapted to it. When adopting these deep learning methods in federated settings, the test accuracy decreases to some extent. For example, the test accuracy of FL-MLP under balanced data case is 3.03 percent smaller than that of MLP approach, this is mainly because of the non-IID nature of datasets generated by the users at different homes. The imbalanced data samples for each user or home will also deteriorate the test accuracy. Nevertheless, our proposed FedHome approach can mitigate the statistical challenge inherent in federated settings (imbalanced and non-IID nature of dataset from different users) and improve the test accuracy to 95.87 percent for balanced data case and 95.41 percent for the imbalanced data case. Specifically, we also examine the performance of FedHome without personalization, which is essentially a federated version of GCAE. The results show that test accuracy will drop by 6.74 and 11.19 percent in balanced and imbalanced data cases, respectively, when compared with FedHome.

As for communication efficiency which is essential in federated learning, we give the number of model parameters of all the FL based approaches in Table 3. Our proposed FedHome framework only has 52,149 model parameters and will reduce the communication load of model transfer in each round significantly compared with FL-MLP whose model parameter number is nearly 30 times more than that of FedHome. Although FL-CNN has fewer parameters than FedHome, it yields 10.8-12.5 percent accuracy degradation. As scaling up CNN size (e.g., width, depth, etc.) is known to be an effective approach for improving model accuracy [41], we also compare our FedHome model with FL-CNN-Large, a CNN-based model whose parameter number is 1.49 times larger than that of FedHome. The experimental results show that FedHome can significantly outperform FL-CNN-Large in both balanced and imbalanced data partitions. Moreover, even FedHome-p, a federated version of GCAE without personalization, can achieve comparable performance to FL-CNN-Large, indicating that GCAE can reduce the communication overhead without sacrificing model accuracy. In short, our experiments demonstrate the effectiveness and applicability of our proposed FedHome framework in realistic computation and communication-constrained in-home health monitoring applications.

To give a more detailed performance evaluation, we also show the average precision, recall and f1-score over all 30 users for each activity. As depicted in Table 4, FedHome achieves excellent performance on all the metrics for all activities. For example, the precision scores for JUM and STD activities are higher than 99 percent, which means that FedHome can accurately identify these types of activities. Even for fall and fall-like activities, the precision and recall scores are all over 90 percent, indicating that our proposed FedHome framework is able to well distinguish falls from similar activities. Thus, for some healthcare applications which focus on fall detection, FedHome can eliminate the risk of fall-induced injuries among the elderly by signaling for help when detecting falls.

TABLE 4
The Average Precision, Recall, and F1-Score for Each Activity

Activity	Precision	Recall	F1-score
Fall	96.24%	90.2%	93.10%
SCH	96.20%	99.54%	97.87%
CSI	93.54%	90.00%	91.30%
CSO	93.40%	93.70%	93.47%
STU	95.23%	92.63%	93.83%
STN	93.60%	93.97%	93.70%
JUM	99.01%	98.83%	98.97%
JOG	96.44%	98.87%	97.63%
STD	99.60%	99.30%	99.50%
WAL	95.40%	98.20%	96.63%

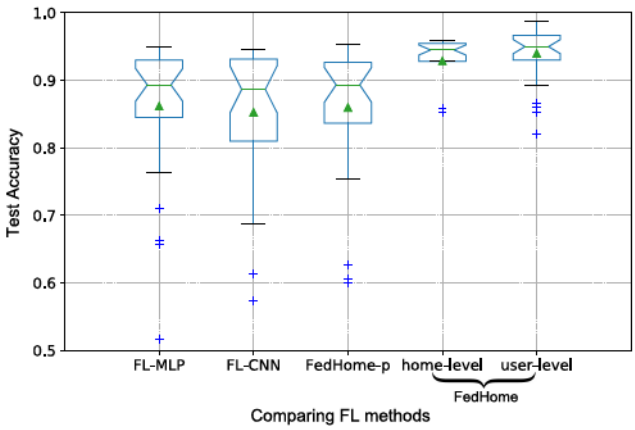


Fig. 8. Test accuracy of different FL approaches under home data partition.

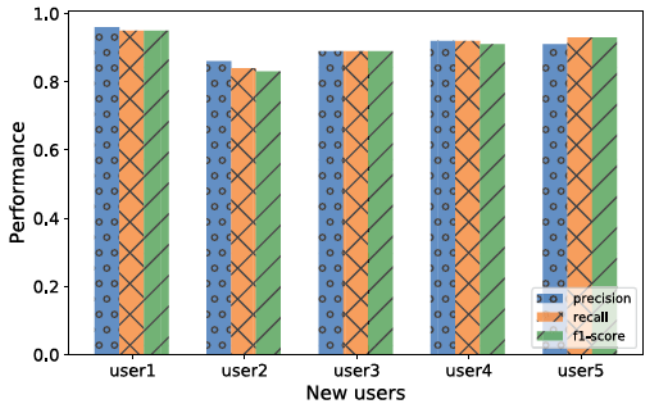


Fig. 9. Performance of FedHome framework when facing new users.

For home data partition, since the training data setting is the same with imbalanced data partition when applied to centralized methods, we only conduct experiments on federated learning approaches. As shown in Fig. 8, all approaches can achieve a high average accuracy of 30 users. However, the test accuracies for these 30 users predicted by FL-CNN vary dramatically. For example, the accuracy of some users may lower than 70 percent while some users can reach a high accuracy of more than 95 percent. By integrating representation ability of AE which can reduce information loss to some extent with parameter sharing, FedHome-p approach can enhance the prediction performance comparing with FL-CNN. Although FL-MLP can achieve more competitive performance than FedHome-p, it is rarely applicable to realistic applications due to its large training parameter size and high communication cost. As family members in the same home trust and share data with each other, they can choose to perform personalization at either home level or user level. For example, users who have insufficient private data to develop accurate local models can choose home-level personalization in order to benefit from their mutual-trusting family members. While users with abundant data samples can conduct user-level personalization with their personal data. The average test accuracies of FedHome personalized at home level and user level are 92.97 and 94.03 percent, respectively. Moreover, we can see that the test accuracies of 30 users vary in a very small scale when performing home-level personalization. This is largely due to the fact that users with highly-skewed personal data can utilize their family members' health data to enhance their own models. While for some users, personalizing their model using their own data can make the refined model more tailored to them.

We also investigate the generalization ability of FedHome when facing new users. We first randomly select 5 users who have not participated in the cloud model training, and then distribute the learned cloud model to these 5 users. After training with only a few rounds using user's personal data, the local model for each user can achieve a good performance. For instance, as depicted in Fig. 9, the precision score of user 1 is 96 percent and the average precision score of all 5 users is 90.8 percent.

5 Related Work

In this section, we elaborate the recent works in in-home health monitoring and federated learning.

5.1 In-Home Health Monitoring

In-home health monitoring improves healthcare effectiveness and lowers healthcare costs by utilizing the mobile and ambient sensors in smart homes, such as smartphones and smartwatches. Due to its importance and urgency in the context of ageing population worldwide, smart healthcare has evoked notable scientific interest [42]. For example, continuous monitoring and automatic classification of tremor severity in Parkinson's Disease are conducted using a wristwatch-type wearable device consisting of an accelerometer and a gyroscope [43]. Physiological signals (e.g., electroencephalography) are used to detect certain diseases, such as seizures [44]. Besides, many studies focus on sensor-based human activity recognition and fall detection due to their wide range of application domains including assisted living, sport, human-computer interaction and healthcare [35], [45]. In these studies, machine learning approaches, such as support vector machine, random forest, are widely used for health status monitoring [34]. With the recent success of deep learning, many healthcare applications adopt some representative networks including convolutional neural network, autoencoder, and have achieved satisfactory performances [35]. It is noteworthy that these traditional health monitoring applications often build models by aggregating all the user data in a centralized location and lack privacy protection and personalization.

Muhammen *et al.* propose a personalized ubiquitous cloud and edge-enabled networked healthcare system called UbeHealth [46]. However, they only emphasize that security policies should be developed and maintained from the device manufacturers, software developing organizations and regulators that check the standard and safety, without protecting user privacy security at the algorithm level [47]. Zhang *et al.* propose a collaborative cloud-edge computation system for personalized driving behavior modeling [48]. Although this method protects user privacy by keeping user data at the edge devices, it requires a large amount of data that are anonymized and integrated from users who would like to share their data in the cloud, which is impractical in real-world applications.

5.2 Federated Learning

Federated learning is first proposed by Google to solve the data island and user privacy problems [49]. The key idea is to build machine learning models based on datasets

distributed across multiple devices while preventing data leakage [50]. McMahan *et al.* introduce the `FederatedAverage` algorithm which has been verified to be robust and has been regarded as the fundamental framework in federated learning settings [17]. While solving the problems of data island and privacy security, federated learning poses new challenges, such as imbalanced and highly-skewed data distribution and high communication cost. Thus, recent improvements have been focusing on overcoming these statistical and communication challenges [49], [51], [52]. Zhao *et al.* propose a data-sharing strategy to improve FedAvg with non-IID by creating a small subset of data which is globally shared among clients [9]. To reduce communication overhead, Jeong *et al.* propose federated distillation, a distributed knowledge distillation method whose communication payload size depends not on the model size but the output dimension [10]. Prior to operating federated distillation, the authors empower each participating device to locally reproduce the data samples of all devices so as to make the training dataset become IID by using a trained generative adversarial network. Nevertheless, this approach requires the user device to upload a few seed data samples to the server, which may put user privacy as risk. There are also some research efforts targeting at making federated learning more personalizable [53]. For example, the proposed FedHealth framework achieves personalized model learning for each organization through knowledge transfer while ignoring the fact that class imbalance problem will bring huge prediction errors [11].

6 Conclusion

In this paper, we propose FedHome, a novel cloud-edge federated learning framework for personalized in-home health monitoring, which achieves privacy protection by keeping user data locally. FedHome aggregates the data from multiple homes and trains a global model without compromising user privacy, and then achieves personalized model learning through knowledge transfer. To tackle the statistical and communication challenge inherent in federated learning, the cloud model to be learned is designed as a generative convolutional autoencoder, which empowers to synthesize samples of minority classes and form a class-balanced dataset in personalization procedure to address the imbalanced and non-IID data dilemma. Moreover, GCAE consists only a small number of model parameters, and thus can significantly reduce the communication overhead during model transfer. Extensive experiments on human activity recognition have demonstrated the effectiveness of the proposed framework in both evaluation performance and communication efficiency. FedHome can be applied to many healthcare applications without incurring data leakage and can be a powerful approach for in-home health monitoring in the future.

Acknowledgments

This work was supported in part by the National Key Research and Development Program of China under grant No.2017YFB1001703; the National Science Foundation of China (No. U20A20159, No. 61972432); the Program for Guangdong Introducing Innovative and Entrepreneurial Teams (No.2017ZT07X355); the Pearl River Talent Recruitment Program (No.2017GC010465).

References

[1] World Health Organization. "Towards policy for health and ageing," Accessed: Dec. 15, 2019. [Online]. Available: http:www.who.int/ageing/publications/alc_fs_ageing_policy.pdf
[2] L. Y. Mano *et al.*, "Exploiting IoT technologies for enhancing health smart homes through patient identification and emotion recognition," *Comput. Commun.*, vol. 89, pp. 178–190, 2016.
[3] P. Verma and S. K. Sood, "Fog assisted-IoT enabled patient health monitoring in smart homes," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1789–1796, Jun. 2018.
[4] S. Hiremath, G. Yang, and K. Mankodiya, "Wearable Internet of Things: Concept, architectural components and promises for person-centered healthcare," in *Proc. 4th Int. Conf. Wireless Mobile Commun. Healthcare-Transforming Healthcare Through Innovations Mobile Wireless Technol.*, 2014, pp. 304–307.
[5] P. Voigt and A. Von dem Bussche, "The EU general data protection regulation (GDPR)," in *A Practical Guide*, 1st Ed., Berlin, Germany: Springer International Publishing, 2017.
[6] N. Inkster, *China's Cyber Power*. Abingdon, U.K.: Routledge, 2018.
[7] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, "Federated learning for healthcare informatics," *J. Healthcare Informat. Res.*, 2020. [Online]. Available: https://doi.org/10.1007/s41666-020-00082-4
[8] W. Y. B. Lim *et al.*, "Federated learning in mobile edge networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 2031–2063, Third Quarter 2020.
[9] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, "Federated learning with non-iid data," 2018, *arXiv: 1806.00582*.
[10] E. Jeong, S. Oh, H. Kim, J. Park, M. Bennis, and S.-L. Kim, "Communication-efficient on-device machine learning: Federated distillation and augmentation under non-IID private data," 2018, *arXiv: 1811.11479*.
[11] Y. Chen, X. Qin, J. Wang, C. Yu, and W. Gao, "FedHealth: A federated transfer learning framework for wearable healthcare," *IEEE Intell. Syst.*, vol. 35, no. 4, pp. 83–93, Jul./Aug. 2020.
[12] S. Silva, B. A. Gutman, E. Romero, P. M. Thompson, A. Altmann, and M. Lorenzi, "Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data," in *Proc. IEEE 16th Int. Symp. Biomed. Imag.*, 2019, pp. 270–274.
[13] Y. Kim, J. Sun, H. Yu, and X. Jiang, "Federated tensor factorization for computational phenotyping," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2017, pp. 887–895.
[14] D. Liu, D. Dligach, and T. Miller, "Two-stage federated phenotyping and patient representation learning," in *Proc. 18th BioNLP Workshop Shared Task*, 2019, pp. 283–291.
[15] Healthcare in Europe, "Federated learning brings ai with privacy to hospitals.," Accessed: Sep. 26, 2020, 2020. [Online]. Available: https://healthcare-in-europe.com/en/home/
[16] Mass General Brigham, "Partners healthcare launches federated learning initiative to deliver AI at the point of care.," Accessed Sep. 26, 2020, 2020. [Online]. Available: https://www.massgeneralbrigham.org/
[17] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th Int. Conf. Artif. Intell. Statist.*, 2017, pp. 1273–1282.
[18] S. Lin, G. Yang, and J. Zhang, "Real-time edge intelligence in the making: A collaborative learning framework via federated meta-learning," 2020, *arXiv: 2001.03229*.
[19] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989.
[20] R. L. Rivest *et al.*, "On data banks and privacy homomorphisms," *Found. Secure Comput.*, vol. 4, no. 11, pp. 169–180, 1978.
[21] L. T. Phong, Y. Aono, T. Hayashi, L. Wang, and S. Moriai, "Privacy-preserving deep learning via additively homomorphic encryption," *IEEE Trans. Inf. Forensics Secur.*, vol. 13, no. 5, pp. 1333–1345, May 2018.

[22] Webank's AI, "Federated ai technology enabler.," Accessed: Sep. 26, 2020, 2020. [Online]. Available: https://www.fedai.org/cn/
[23] P. Kairouz *et al.*, "Advances and open problems in federated learning," 2019, *arXiv: 1912.04977*.
[24] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, pp. 3371–3408, 2010.
[25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.
[26] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.
[27] D. Elreedy and A. F. Atiya, "A comprehensive analysis of synthetic minority oversampling technique (smote) for handling class imbalance," *Inf. Sci.*, vol. 505, pp. 32–64, 2019.
[28] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, 2014, pp. 3320–3328.
[29] S.-A. Rebuffi, A. Kolesnikov, G. Sperl, and C. H. Lampert, "iCaRL: Incremental classifier and representation learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2001–2010.
[30] W. Jung, A. W. Mulyadi, and H.-I. Suk, "Unified modeling of imputation, forecasting, and prediction for ad progression," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2019, pp. 168–176.
[31] C. Lian, M. Liu, L. Wang, and D. Shen, "End-to-end dementia status prediction from brain MRI using multi-task weakly-supervised attention network," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2019, pp. 158–167.
[32] G. Vavoulas, C. Chatzaki, T. Malliotakis, M. Pediaditis, and M. Tsiknakis, "The MobiAct dataset: Recognition of activities of daily living using smartphones," in *Proc. 2nd Int. Conf. Inf. Commun. Technol. Ageing Well e-Health*, 2016, pp. 143–151.
[33] E. S. Grood and W. J. Suntay, "A joint coordinate system for the clinical description of three-dimensional motions: Application to the knee," *J. Biomech. Eng. Trans. ASME*, vol. 105, no. 2, pp. 136–144, 1983.
[34] J. A. Ward, G. Pirkl, P. Hevesi, and P. Lukowicz, "Towards recognising collaborative activities using multiple on-body sensors," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.: Adjunct*, 2016, pp. 221–224.
[35] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A survey," *Pattern Recognit. Lett.*, vol. 119, pp. 3–11, 2019.
[36] G. Shi, C. S. Chan, W. J. Li, K.-S. Leung, Y. Zou, and Y. Jin, "Mobile human airbag system for fall protection using MEMS sensors and embedded SVM classifier," *IEEE Sensors J.*, vol. 9, no. 5, pp. 495–503, May 2009.
[37] J. He, C. Hu, and X. Wang, "A smart device enabled system for autonomous fall detection and alert," *Int. J. Distrib. Sensor Netw.*, vol. 12, no. 2, 2016, Art. no. 2308183.
[38] J. Yuan, K. K. Tan, T. H. Lee, and G. C. H. Koh, "Power-efficient interrupt-driven algorithms for fall detection and classification of activities of daily living," *IEEE Sensors J.*, vol. 15, no. 3, pp. 1377–1387, Mar. 2015.
[39] J. He, Z. Zhang, X. Wang, and S. Yang, "A low power fall sensing technology based on FD-CNN," *IEEE Sensors J.*, vol. 19, no. 13, pp. 5110–5118, Jul. 2019.
[40] M. Chen, X. Shi, Y. Zhang, D. Wu, and M. Guizani, "Deep features learning for medical image analysis with convolutional autoencoder neural network," *IEEE Trans. Big Data*, early access, Jun. 20, 2017, doi: 10.1109/TBDATA.2017.2717439.
[41] C. He, S. Avestimehr, and M. Annavaram, "Group knowledge transfer: Collaborative training of large CNNs on the edge," 2020, *arXiv: 2007.14513*.
[42] D. J. Cook, G. Duncan, G. Sprint, and R. L. Fritz, "Using smart city technology to make healthcare smarter," *Proc. IEEE*, vol. 106, no. 4, pp. 708–722, Apr. 2018.
[43] H. Jeon *et al.*, "Automatic classification of tremor severity in parkinson's disease using a wearable device," *Sensors*, vol. 17, no. 9, 2017, Art. no. 2067.
[44] M. E. L. Menshawy, A. Benharref, and M. Serhani, "An automatic mobile-health based approach for EEG epileptic seizures detection," *Expert Syst. Appl.*, vol. 42, no. 20, pp. 7157–7174, 2015.
[45] B. Chikhaoui, F. Gouineau, and M. Sotir, "A CNN based transfer learning model for automatic activity recognition from accelerometer sensors," in *Proc. Int. Conf. Mach. Learn. Data Mining Pattern Recognit.*, 2018, pp. 302–315.
[46] T. Muhammed, R. Mehmood, A. Albeshri, and I. Katib, "UbeHealth: A personalized ubiquitous cloud and edge-enabled networked healthcare system for smart cities," *IEEE Access*, vol. 6, pp. 32258–32285, 2018.
[47] J. Sametinger, J. W. Rozenblit, R. L. Lysecky, and P. Ott, "Security challenges for medical devices," *Commun. ACM*, vol. 58, no. 4, pp. 74–82, 2015.
[48] X. Zhang, M. Qiao, L. Liu, Y. Xu, and W. Shi, "Collaborative cloud-edge computation for personalized driving behavior modeling," in *Proc. 4th ACM/IEEE Symp. Edge Comput.*, 2019, pp. 209–221.
[49] J. Konečnỳ, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," 2016, *arXiv:1610.05492*.
[50] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, 2019, Art. no. 12.
[51] V. Smith, C.-K. Chiang, M. Sanjabi, and A. S. Talwalkar, "Federated multi-task learning," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 4424–4434.
[52] Y. Lin, S. Han, H. Mao, Y. Wang, and W. J. Dally, "Deep gradient compression: Reducing the communication bandwidth for distributed training," in *Proc. 6th Int. Conf. Learn. Representations*, 2018.
[53] F. Chen, M. Luo, Z. Dong, Z. Li, and X. He, "Federated meta-learning with fast convergence and efficient communication," 2018, *arXiv: 1802.07876*.

Qiong Wu received the BS and ME degrees from the School of Data and Computer Science, Sun Yat-sen University (SYSU), Guangzhou, China, in 2017 and 2019, respectively. She is currently working toward the PhD degree in the School of Data and Computer Science, Sun Yat-sen University, China. Her primary research interests include social data analysis, mobile edge computing, and federated learning.

Xu Chen (Senior Member, IEEE) received the PhD degree in information engineering from the Chinese University of Hong Kong, Hong Kong, in 2012, and worked as a postdoctoral research associate at Arizona State University, Tempe, Arizona, from 2012 to 2014, and a Humboldt Scholar fellow at the Institute of Computer Science of the University of Goettingen, Germany from 2014 to 2016. He is a full professor with Sun Yat-sen University, Guangzhou, China, and the vice director of National and Local Joint Engineering Laboratory of Digital Home Interactive Applications. He received the prestigious Humboldt research fellowship awarded by the Alexander von Humboldt Foundation of Germany, 2014 Hong Kong Young Scientist Runner-up Award, 2017 IEEE Communication Society Asia-Pacific Outstanding Young Researcher Award, 2017 IEEE ComSoc Young Professional Best Paper Award, Honorable Mention Award of 2010 IEEE International Conference on Intelligence and Security Informatics (ISI), Best Paper Runner-up Award of 2014 IEEE International Conference on Computer Communications (INFOCOM), and Best Paper Award of 2017 IEEE Intranational Conference on Communications (ICC). He is currently an area editor of the *IEEE Open Journal of the Communications Society*, an associate editor of the *IEEE Transactions Wireless Communications*, the *IEEE Internet of Things Journal* and the *IEEE Journal on Selected Areas in Communications (JSAC)* Series on Network Softwarization and Enablers.

Zhi Zhou (Member, IEEE) received the BS, ME, and PhD degrees all from the School of Computer Science and Technology, Huazhong University of Science and Technology (HUST), Wuhan, China, in 2012, 2014, and 2017, respectively. He is currently an associate professor in the School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China. In 2016, he was a visiting scholar at the University of Göttingen, Germany. He was nominated for the 2019 China Computer Federation CCF Outstanding Doctoral Dissertation Award, the sole recipient of the 2018 ACM Wuhan & Hubei Computer Society Doctoral Dissertation Award, and a recipient of the Best Paper Award of IEEE UIC 2018. His research interests include edge computing, cloud computing, and distributed systems.

Junshan Zhang (Fellow, IEEE) received the PhD degree from the School of ECE, Purdue University, West Lafayette, Indiana, in 2000. He joined the School of ECEE, Arizona State University, Tempe, Arizona, in August 2000, where he has been the Fulton chair professor since 2015. His research interests are in the general field of information networks and data science, including communication networks, Internet of Things (IoT), fog computing, social networks, and smart grid. He is a a recipient of the ONR Young Investigator Award, in 2005 and the NSF CAREER award, in 2003. He received the IEEE Wireless Communication Technical Committee Recognition Award, in 2016. His papers have won several awards, including the Kenneth C. Sevcik Outstanding Student Paper Award of ACM SIGMETRICS/IFIP Performance 2016, the Best Paper Runner-up Award of IEEE INFOCOM 2009 and IEEE INFOCOM 2014, and the Best Paper Award at IEEE ICC 2008 and ICC 2017. He was TPC co-chair for a number of major conferences in communication networks, including IEEE INFOCOM 2012 and ACM MOBIHOC 2015. He was the general chair for ACM/ IEEE SEC 2017, WiOPT 2016, and IEEE Communication Theory Workshop 2007. He was a distinguished lecturer of the IEEE Communications Society. He was an associate editor for the *IEEE Transactions on Wireless Communications*, an editor for the *Computer Networks Journal*, and an editor for the *IEEE Wireless Communication Magazine*. He is currently serving as the editor-in-chief for the *IEEE Transactions on Wireless Communications*, an editor-at-large for the *IEEE/ACM Transactions on Networking*, and an editor for the *IEEE Network*.