FedHealth: A Federated Transfer Learning Framework for Wearable Healthcare

Yiqiang Chen[1,2,3*], **Jindong Wang**[4], **Chaohui Yu**[1,2], **Wen Gao**[3], **Xin Qin**[1,2]
[1]Beijing Key Lab. of Mobile Computing and Pervasive Devices, Inst. of Computing Tech., CAS
[2]University of Chinese Academy of Sciences, Beijing, China
[3]Pengcheng Laboratory, Shenzhen, China
[4]Microsoft Research Asia, Beijing, China
yqchen@ict.ac.cn, jindong.wang@microsoft.com

Abstract

With the rapid development of computing technology, wearable devices such as smart phones and wristbands make it easy to get access to people's health information including activities, sleep, sports, etc. Smart healthcare achieves great success by training machine learning models on large quantity of user data. However, there are two critical challenges. Firstly, user data often exists in the form of isolated islands, making it difficult to perform aggregation without compromising privacy security. Secondly, the models trained on the cloud fail on personalization. In this paper, we propose FedHealth, the first federated transfer learning framework for wearable healthcare to tackle these challenges. FedHealth performs data aggregation through federated learning, and then builds personalized models by transfer learning. It is able to achieve accurate and personalized healthcare without compromising privacy and security. Experiments demonstrate that FedHealth produces higher accuracy (5.3% improvement) for wearable activity recognition when compared to traditional methods. FedHealth is general and extensible and has the potential to be used in many healthcare applications.

1 Introduction

Activities of daily living (ADL) are highly related to people's health. Recently, the development of wearable technologies helps people to understand their health status by tracking activities using wearable devices such as smartphone, wristband, and smart glasses. Wearable healthcare has the potential to provide early warnings to several cognitive diseases such as Parkinson's [Chen *et al.*, 2017; Chen *et al.*, 2019] and small vessel diseases [Chen *et al.*, 2018b]. Other applications include mental health assessment [Wang *et al.*, 2014], fall detection [Wang *et al.*, 2017b], and sports monitoring [Wang *et al.*, 2019a]. In fact, there is a growing trend for wearable healthcare over the years [Andreu-Perez *et al.*, 2015; Hiremath *et al.*, 2014].

*Corresponding Author

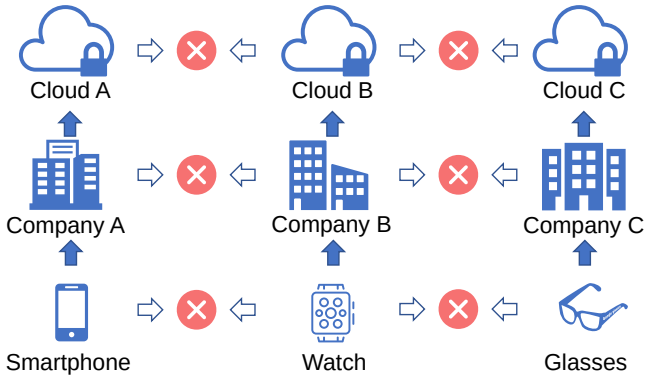

Figure 1: The data islanding and personalization problems in wearable healthcare

In healthcare applications, machine learning models are often trained on sufficient user data to track health status. Traditional machine learning approaches such as Support Vector Machines (SVM), Decision Tree (DT), and Hidden Markov Models (HMM) are adopted in many healthcare applications [Ward *et al.*, 2016]. The recent success of deep learning achieves satisfactory performances by training on larger sizes of user data. Representative networks include Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Autoencoders [Wang *et al.*, 2019a].

Unfortunately, there are two critical challenges in today's wearable healthcare (Figure 1). First of all, in real life, data often exists in the form of isolated islands. Although there are plenty of data in different organizations, institutes, and subjects, it is not possible to share them due to privacy and security concerns. In Figure 1, when the same user uses different products from two companies, his data stored in two clouds cannot be exchanged. This makes it hard to train powerful models using these valuable data. Additionally, recently, China, the United States, and the European Union enforced the protection of user data via different regularizations [Inkster, 2018; Voigt and Von dem Bussche, 2017]. Hence, the acquisition of massive user data is not possible in real applications.

The other important issue is personalization. Most of the methods are based on a common server model for nearly all users. After acquiring sufficient user data to train a satis-

factory machine learning model, the model itself is then distributed to all the user devices on which the daily health information can be tracked. This process lacks personalization. As can be seen, different users have different physical characteristics and daily activity patterns. Therefore, the common model fails to perform personalized healthcare.

In this paper, we propose *FedHealth*, the first federated transfer learning framework for wearable healthcare. FedHealth can solve both of the data islanding and personalization problems. Through federated learning [Yang *et al.*, 2019; Yang *et al.*, 2018] and homomorphic encryption [Rivest *et al.*, 1978], FedHealth aggregates the data from separate organizations to build powerful machine learning models with the users' privacy well preserved. After the cloud model is built, FedHealth utilizes transfer learning [Pan and Yang, 2010] methods to achieve personalized model learning for each organization. The framework can incrementally update. FedHealth is extensible and can be deployed to many healthcare applications to continuously enhance their learning abilities in real life.

In summary, this paper makes the following contributions:

1. We propose FedHealth, the first federated transfer learning framework for wearable healthcare, which aggregates the data from different organizations without compromising privacy security, and achieves personalized model learning through knowledge transfer.

2. We show the excellent performance achieved by FedHealth in smartphone based human activity recognition. Experiments show that FedHealth dramatically improves the recognition accuracy by 5.3% compared to traditional learning approaches.

3. FedHealth is extensible and can be the standard framework to many healthcare applications. With the users' privacy well preserved and good performance achieved, it can be easily deployed to other healthcare applications.

2 Related Work

In this section, we introduce the related work in three aspects: wearable healthcare, federated machine learning, and transfer learning.

2.1 Wearable Healthcare

Certain activities in daily life reflect early signals of some cognitive diseases [Atkinson *et al.*, 2007; Michalak *et al.*, 2009]. For instance, the change of gait may result in small vessel disease or stroke. A lot of researchers pay attention to monitor users' activities using body-worn sensors [Voigt and Von dem Bussche, 2017], through which daily activities and sports activities can be recognized. With the development of wearable technology, smartphone, wristbands, and smart glasses provide easy access to this information. Many endeavors have been made [Wang *et al.*, 2014; Albinali *et al.*, 2010; Wang *et al.*, 2017b]. Other than activities, physiological signals can also help to detect certain diseases. EEG (electroencephalography) is used to detect seizures [Menshawy *et al.*, 2015; Hiremath *et al.*, 2014]. Authors can also use RGB-D cameras to detect users' activities [Lei *et al.*, 2012; Rashidi and Cook, 2010]. For a complete survey on sensors based activity recognition and healthcare, interested readers are recommended to refer to [Wang *et al.*, 2019a].

It is noteworthy that traditional healthcare applications often build the model by aggregating all the user data. However, in real applications, data are often separate and cannot be easily shared due to privacy issues [Inkster, 2018; Voigt and Von dem Bussche, 2017]. Moreover, the models built by applications lack the ability of personalization.

2.2 Federated Machine Learning

A comprehensive survey on federated learning is in [Yang *et al.*, 2019]. Federated machine learning was firstly proposed by Google [Konečnỳ *et al.*, 2016; Konečnỳ *et al.*, 2016], where they trained machine learning models based on distributed mobile phones all over the world. The key idea is to protect user data during the process. Since then, other researchers started to focus on privacy-preserving machine learning [Bonawitz *et al.*, 2017; Shokri and Shmatikov, 2015; Geyer *et al.*, 2017], federated multi-task learning [Smith *et al.*, 2017], as well as personalized federated learning [Chen *et al.*, 2018a]. Federated learning has the ability to resolve the data islanding problems by privacy-preserving model training in the network.

According to [Yang *et al.*, 2019], federated learning can mainly be classified into three types: 1) horizontal federated learning, where organizations share partial features; 2) vertical federated learning, where organizations share partial samples; and 3) federated transfer learning, where neither samples or features have much in common. FedHealth belongs to federated transfer learning category. It is the first of its kind tailored for wearable healthcare applications.

2.3 Transfer Learning

Transfer learning aims at transferring knowledge from existing domains to a new domain. In the setting of transfer learning, the domains are often different but related, which makes knowledge transfer possible. The key idea is to reduce the distribution divergence between different domains. To this end, there are mainly two kinds of approaches: 1) instance reweighting [Huang *et al.*, 2012; Huang *et al.*, 2007], which reuses samples from the source domain according to some weighting technique; and 2) feature matching, which either performs subspace learning by exploiting the subspace geometrical structure [Wang *et al.*, 2018b; Sun *et al.*, 2016; Fernando *et al.*, 2013; Gong *et al.*, 2012], or distribution alignment to reduce the marginal or conditional distribution divergence between domains [Wang *et al.*, 2019b; Wang *et al.*, 2018a; Wang *et al.*, 2017a; Pan *et al.*, 2011; **?**]. Recently, deep transfer learning methods have made considerable success in many application fields [Rozantsev *et al.*, 2019; Ganin and Lempitsky, 2015; Tzeng *et al.*, 2014]. For a complete survey, please refer to [Pan and Yang, 2010].

FedHealth is mainly related to deep transfer learning. Most of the methods assume the availability of training data, which is not realistic. FedHealth makes it possible to do deep transfer learning in the federated learning framework without accessing the raw user data. Therefore, it is more secure.

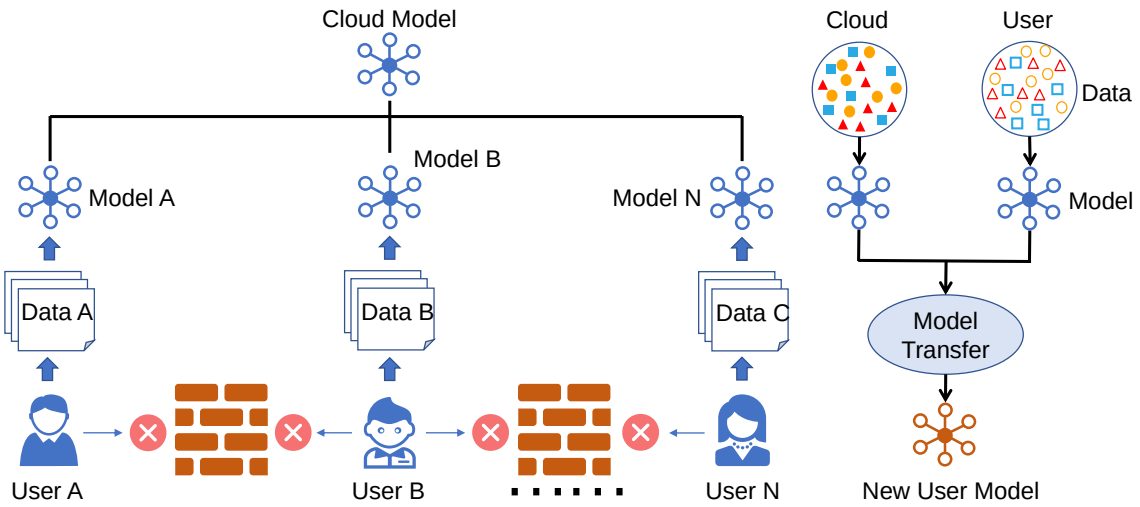

Figure 2: Overview of the FedHealth framework. "User" represents organizations

3 The Proposed FedHealth Framework

In this section, we introduce the FedHealth framework for federated transfer learning based wearable healthcare.

3.1 Problem Definition

We are given data from N different users (organizations), denoted the users by $\{\mathcal{S}_1, \mathcal{S}_2, \cdots, \mathcal{S}_N\}$ and the sensor readings they provide are denoted by $\{\mathcal{D}_1, \mathcal{D}_2, \cdots, \mathcal{D}_N\}$. Conventional methods train a model $\mathcal{M}_{ALL}$ by combining all the data $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2 \cup \cdots \cup \mathcal{D}_N$. All the data have different distributions. In our problem, we want to collaborate all the data to train a federated model $\mathcal{M}_{FED}$, where any user $\mathcal{S}_i$ does not expose its data $\mathcal{D}_i$ to each other. If we denote the accuracy as $\mathcal{A}$, then the objective of FedHealth is to ensure the accuracy of federated learning is close to or superior to that of conventional learning denoted by:

$$\mathcal{A}_{FED} - \mathcal{A}_{ALL} > \Delta, \tag{1}$$

where Δ is an extremely small non-negative real number.

3.2 Overview of the Framework

FedHealth aims to achieve accurate personal healthcare through federated transfer learning without compromising privacy security. Figure 2 gives an overview of the framework. Without loss of generality, we assume there are 3 users (organizations) and 1 server, which can be extended to the more general case. The framework mainly consists of four procedures. First of all, the cloud model on the server end is train based on public datasets. Then, the cloud model is distributed to all users where each of them can train their own model on their data. Subsequently, the user model can be uploaded to the cloud to help train a new cloud model. Note that this step does not share any user data or information but the encrypted model parameters. Finally, each user can train personalized models by integrating the cloud model and its previous model and data for personalization. In this step, since there is large distribution divergence between cloud and user model, transfer learning is performed to make the model more tailored to the user (right part in Figure 2). It is noteworthy that all the parameter sharing processes does not involve any leakage of user data. Instead, they are finished through homomorphic encryption [Rivest *et al.*, 1978].

The federated learning paradigm is the main computing model for the whole FedHealth framework. It deals with model building and parameter sharing during the entire process. After the server model is learned, it can be directly applied to the user. This is just what traditional healthcare applications do for model learning. It is obvious that the samples in the server are having highly different probability distribution with the data generated by each user. Therefore, the common model fails in personalization. Additionally, user models cannot easily be updated continuously due to the privacy security issue.

3.3 Federated Learning

FedHealth adopts the federated learning paradigm [Yang *et al.*, 2019] to achieve encrypted model training and sharing. This step mainly consists of two critical parts: cloud and user model learning. After obtaining the server model, it is distributed to the user end to help them train their own models. As for each user, it trains its own model with the help of the server model.

In FedHealth, we adopt deep neural networks to learn both the cloud and user model. Deep networks perform end-to-end feature learning and classifier training by taking the raw inputs of the user data as inputs. Let f_S denote the server model to be learned, then the learning objective becomes:

$$\arg\min_{\Theta} \mathcal{L} = \sum_{i=1}^{n} \ell(y_i, f_S(\mathbf{x}_i)), \tag{2}$$

where $\ell(\cdot, \cdot)$ denotes the loss for the network, e.g. cross-entropy loss for classification tasks. $\{\mathbf{x}_i, y_i\}_{i=1}^{n}$ are samples from the server data with n their sizes. Θ denotes all the parameters to be learned, i.e. the weight and bias.

After acquiring the cloud model, it is distributed to all the users. As we can see from the "wall" in Figure 2, direct shar-

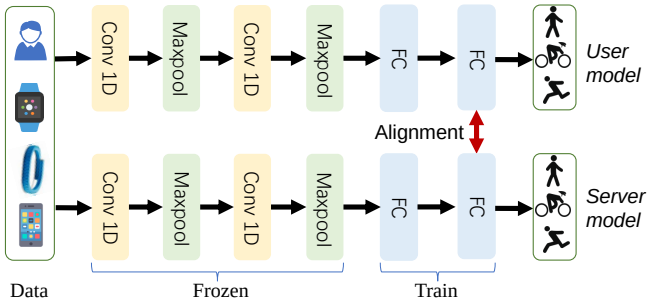

Figure 3: The transfer learning process of FedHealth

ing of user information is forbidden. This process uses homomorphic encryption [Rivest *et al.*, 1978] to avoid information leakage. Since the encryption is not our main contribution, we will show the process of additively homomorphic encryption using real numbers. The encryption scheme of the weight matrix and bias vector are following the same idea. The additively homomorphic encryption of a real number a is denoted as $\langle a \rangle$. In additively homomorphic encryption, for any two numbers a and b, we have $\langle a \rangle + \langle b \rangle = \langle a + b \rangle$. Therefore, the parameter sharing can be done without leaking any information from the users. Through federated learning, we can aggregate user data without compromising privacy security.

Technically, the learning objective for user u is denoted as:

$$\arg\min_{\Theta^u} \mathcal{L}_1 = \sum_{i=1}^{n^u} \ell(y_i^u, f_u(\mathbf{x}_i^u)). \tag{3}$$

It is important to note that FedHealth does not perform parameter sharing as in [Cheng *et al.*, 2019] for computational efficiency. After all the user model f_u is trained, it is uploaded to the server for aggregation. As for aggregation, server can align the old model with the model from each user subsequently. Considering the computational burden, server can also achieve scheduled update (e.g. every night) using uploading user models. The result is a new server model f_S'. Note that the new server model f_S' is based on the knowledge from all users. Therefore, it has better generalization ability.

3.4 Transfer Learning

Federated learning solves the data islanding problem. Therefore, we can build models using all the user data. Another important factor is the personalization. Even if we can directly use the cloud model, it still performs poor on a particular user. This is due to the distribution difference between the user and the cloud data. The common model in the server only learns the coarse features from all users, while it fails in learning the fine-grained information on a particular user.

In this paper, FedHealth uses transfer learning to build a personalized model for each user. Recall that features in deep networks are highly transferable in the lower levels of the network since they focus on learning common and low-level features. The higher layers learn more specific features to the task [Yosinski *et al.*, 2014]. In this way, after obtaining the parameters of the cloud model, we can perform transfer learning on the user to learn their personalized models.

Figure 3 presents the process of transfer learning for a specific convolutional neural network (CNN). Suppose the network is composed of two convolutions layers (`conv1, conv2`), two max-pooling layers (`pool1, pool2`), two fully connected layers (`fc1, fc2`), and one softmax layer for classification. The network is designed for human activity recognition where the input data is the activity signals for a user and the output is his activity classes.

Algorithm 1 The learning procedure of FedHealth

Input: Data from different users $\{\mathcal{D}_1, \mathcal{D}_2, \cdots, \mathcal{D}_N\}$, η
Output: Personalized user model f_u

1: Construct a cloud model f_S using Eq. (2)
2: Distribute f_S to all users via homomorphic encryption
3: Train user models using Eq. (3)
4: Update all user models to the server using homomorphic encryption. Then server update its model by aligning with user model
5: Distribute f_S' to all users, then perform transfer learning on each user to get their personalized model f_u using Eq. (6)
6: Repeat the above procedures with the continuously emerging user data

In model transfer, we think that the convolution layers aims at extracting low-level features about activity recognition. Thus we keep these layers along with the max-pooling layers frozen, which means we do not update their parameters in backpropagation. As for the fully connected layers `fc1` and `fc2`, since they are higher level, we believe they focus on learning specific features for the task and user. Therefore, we update their parameters during training. The softmax serves as the classification function, which can be formulated as:

$$y_j = \frac{e^{z_c}}{\sum_{c=1}^{C} e^{z_c}}, \tag{4}$$

where z_c denotes the learned probability for class C, and y_j is the final classification result.

FedHealth adapts the inputs from different domains by replacing `fc2` with an `alignment` layer. This is strictly different that in DDC [Tzeng *et al.*, 2014] and other recent methods where we have access to both the source and target data. In our problem, we only have the user data and the cloud model. To this end, we borrow the idea from [Rozantsev *et al.*, 2018] and regularize the weights. Given the network from the server and user, we add a correlation alignment [Sun *et al.*, 2016] layer before the softmax layer to further adapt the domains. This alignment function is used to align the second-order statistics between the inputs. Formally, the loss of correlation alignment is computed as follows:

$$\ell_{CORAL} = \frac{1}{4d^2} \|C_S - C_T\|_F^2 \tag{5}$$

where $\|\cdot\|_F^2$ denotes the squared matrix Frobenius norm and d is the dimension of the embedding features. C_S and C_T are the covariance matrices of the source and target weights computed by [Sun *et al.*, 2016]. Therefore, denote η the trade-off parameter, the loss for the user model is computed by:

$$\arg\min_{\Theta_u} \mathcal{L}_u = \sum_{i=1}^{n_u} \ell(y_i^u, f_u(\mathbf{x}_i^u)) + \eta \ell_{CORAL}. \tag{6}$$

3.5 Learning Process

The learning procedure of FedHealth is presented in Algorithm 1. Note that this framework works continuously with the new emerging user data. FedHealth can update the user model and cloud model simultaneously when facing new user data. Therefore, the longer the user uses the product, the more personalized the model can be. Other than transfer learning, FedHealth can also embed other popular methods for personalization such as incremental learning [Rebuffi *et al.*, 2017].

The entire framework can also adopt other machine learning methods other than deep networks. For instance, the gradient boosting decision tree can be integrated into the framework to harness the power of ensemble learning. These lightweight models can be deployed to computation restricted wearable devices. This makes FedHealth more general to real applications.

4 Experiments

In this section, we evaluate the performance of the proposed FedHealth framework via extensive experiments on human activity recognition.

4.1 Datasets

We adopt a public human activity recognition dataset called UCI Smartphone [Anguita *et al.*, 2012]. This dataset contains 6 activities collected from 30 users. The 6 activities are WALKING, WALKING-UPSTAIRS, WALKING-DOWNSTAIRS, SITTING, STANDING, and LAYING. There are 30 volunteers within an age bracket of 19-48 years. Each volunteer wears a smartphone (Samsung Galaxy S II) on the waist. Using its embedded accelerometer and gyroscope, collectors captured 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz. The experiments have been video-recorded to label the data manually. The obtained dataset has been randomly partitioned into two sets, where 70% of the volunteers were selected for generating the training data and 30% the test data. There are 10,299 instances in total. The statistical information of the dataset is shown in Table 1.

Table 1: Statistical information of the dataset

Subject	Activity	Sampling rate	Sensor	Instance	Channel
30	6	50 Hz	Accelerometer Gyroscope	10,299	9

In order to construct the problem situation in FedHealth, we change the standard setting for the dataset. We extracted 5 subjects (Subject IDS $26 \sim 30$) and regard them as the isolated users which cannot share data due to privacy security. Data on the remaining 25 users are used to train the cloud model. Henceforth, the objective is to use the cloud model and all the 5 isolated subjects to improve the activity recognition accuracy on the 5 subjects without compromising the privacy. In short, it is a variant of the framework in Figure 2 where there are 5 users.

Table 2: Classification accuracy (%) of the test subject

Subject	KNN	SVM	RF	NoFed	FedHealth
P1	83.8	81.9	87.5	94.5	**98.8**
P2	86.5	96.9	93.3	94.5	**98.8**
P3	92.2	97.2	88.9	93.4	**100.0**
P4	83.1	95.9	91.0	95.5	**99.4**
P5	90.5	98.6	91.6	92.6	**100.0**
AVG	87.2	94.1	90.5	94.1	**99.4**

4.2 Implementation Details

On both the server and the user end, we adopt a CNN for training and prediction. The network is composed of 2 convolutional layers, 2 pooling layers, and 3 fully connected layers. The network adopts a convolution size of 1×9. It uses mini-batch Stochastic Gradient Descent (SGD) for optimization. During training, we use 70% of the training data for model training, while the rest 30% is for model evaluation. We fix $\eta = 0.01$. We set the learning rate to be 0.01 with batch size of 64 and training epochs fixed to 80. The accuracy of user u is computed as $\mathcal{A}_u = \frac{|\mathbf{x}:\mathbf{x}\in\mathcal{D}_u \wedge \hat{y}(\mathbf{x})=y(\mathbf{x})|}{|\mathbf{x}:\mathbf{x}\in\mathcal{D}_u|}$, where $y(\mathbf{x})$ and $\hat{y}(\mathbf{x})$ denote the true and predicted labels on sample $\mathbf{x}$, respectively.

We follow [Rivest *et al.*, 1978] for homomorphic encryption in federated learning. During transfer learning, we freeze all the convolutional and pooling layers in the network. Only the parameters of the fully connected layers are updated with SGD. To show the effectiveness of FedHealth, we compare its performance with traditional learning, where we record the performances on each subject using the server model only. For notational brevity, we use `NoFed` to denote this setting. We also compare the performances of KNN, SVM, and random forest (RF) with FedHealth. The hyperparameters of all the comparison methods are tuned using cross-validation. For the fair study, we run all the experiments 5 times to record the average accuracies.

4.3 Classification Accuracy

The classification accuracies of activity recognition for each subject are shown in Table 2. The results indicate that our proposed FedHealth achieves the best classification accuracy on all users. Compared to NoFed, it significantly improves the average results by **5.3%**. Compared to the traditional methods (KNN, SVM, and RF), FedHealth also greatly improves the recognition results. In short, it demonstrates the effectiveness of our proposed FedHealth framework.

The results also show that for activity recognition, the deep methods (NoFed and FedHealth) achieve better results than traditional methods. This is due to the representation capability of deep neural networks, while traditional methods have to rely on hand-crafted feature learning. Another advantage of deep learning is that the models can be updated online and incrementally without retraining, while traditional methods require further incremental algorithms. This property is extremely valuable in federated transfer learning where model reuse is important and helpful.

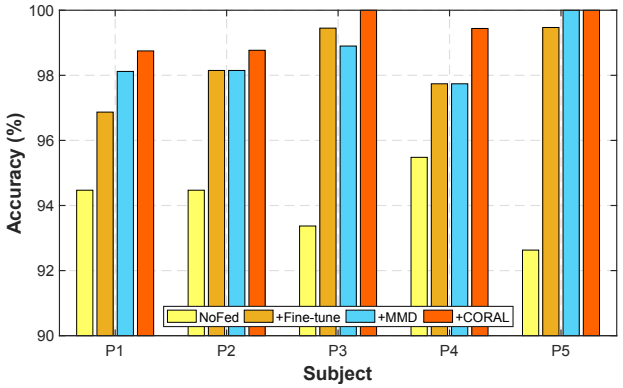

Figure 4: Extending FedHealth with other transfer learning methods

4.4 Evaluation of Extensibility

In this section, we analyze the extensibility of FedHealth with different transfer learning approaches. We compare its performance with two methods: 1) fine-tuning, which only fine-tunes the network on each subject without explicitly reducing the distribution divergence between domains; and 2) transfer with MMD (Maximum Mean Discrepancy) [Wang *et al.*, 2018b], which replaces the alignment loss with MMD loss. The comparison results are shown in Figure 4.

From the results, we can see that other than the alignment loss, FedHealth can also achieve promising results using fine-tuning or MMD. The results of transfer learning significantly outperform no transfer by 4% on average accuracy. This indicates that the transfer learning procedure of FedHealth is highly effective and extensible. Therefore, FedHealth is general and can be extended in many applications by integrating other transfer learning algorithms. Moreover, the federated learning procedure can also be extended using other encryption methods, which can be the future research.

4.5 Detailed Analysis

We provide detailed analysis to FedHealth via comparing its confusion matrix with that of NoFed. The confusion matrix is known as an effective metric to show the efficacy of a method since it provides fine-grained classification results on each task. For simplicity, we show the confusion matrices of subject 2 in Table 3. The results of other subjects follow the same tendency. Along with the confusion matrix, the precision (P), recall (R), and macro $F1$ score ($F1$) are all computed to give a thorough view of the results.

Combining the results in Table 2 and 3, we can clearly see that FedHealth can not only achieve the best accuracy, but also reach the best precision, recall, and $F1$ scores. The confusion matrix shows that FedHealth can reduce the misclassification rate, especially on class $C1$ (Walking). Since walking is the most common activities in healthcare, it means that FedHealth is effective in recognizing this activity. To summarize, FedHealth is more accurate in recognizing personalized activities, which makes it more advantageous in healthcare applications.

Table 3: Classification report of NoFed and FedHealth

NoFed

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	P	R	$F1$
$C1$	70.4%	7.4%	22.2%				1	0.7	0.83
$C2$		96.9%	3.1%				0.94	0.97	0.95
$C3$			100%				0.74	1	0.85
$C4$				100%			1	1	1
$C5$					100%		1	1	1
$C6$						100%	1	1	1
			Average				0.96	0.94	0.94

FedHealth

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	P	R	$F1$
$C1$	**88.9%**	3.7%	7.4%				1	**0.89**	**0.94**
$C2$		**100%**					**0.97**	1	**0.98**
$C3$			100%				**0.91**	1	**0.95**
$C4$				100%			1	1	1
$C5$					100%		1	1	1
$C6$						100%	1	1	1
			Average				**0.98**	**0.98**	**0.98**

5 Discussions

FedHealth is a general framework for wearable healthcare. This paper provides a specific implementation and evaluation of this idea. It is adaptable to several healthcare applications. In this section, we discuss its potential to be extended and deployed to other situations with possible solutions.

1. FedHealth with incremental learning. Incremental learning [Rebuffi *et al.*, 2017] has the ability to update the model with the gradually changing time, environment, and users. In contrast to transfer learning that focuses on model adaptation, incremental learning makes it possible to update the model in real-time without much computation.

2. FedHealth as the standard for wearable healthcare in the future. FedHealth provides such a platform where all the companies can safely share data and train models. In the future, we expect that FedHealth be implemeted with blockchain technology [Zheng *et al.*, 2018] where user data can be more securely stored and protected. We hope that FedHealth can become the standard for wearable healthcare.

3. FedHealth to be applied in more applications. This work mainly focuses on the possibility of federated transfer learning in healthcare via activity recognition. In real situations, FedHealth can be deployed at large-scale to more healthcare applications such as elderly care, fall detection, cognitive disease detection, etc. We hope that through FedHealth, federated learning can become federated computing which can become a new computing model in the future.

6 Conclusions and Future Work

In this paper, we propose FedHealth, the first federated transfer learning framework for wearable healthcare. FedHealth aggregates the data from different organizations without compromising privacy security, and achieves personalized model learning through knowledge transfer. Experiments on human activity recognition have demonstrated the effectiveness of the framework. We also present a detailed discussion for its potential from specific technical improvement to the potential for healthcare applications.

FedHealth opens a new door for future research in wearable healthcare. In the future, we plan to extend FedHealth to the detection of Parkinson's disease where it can be deployed in hospitals.

Acknowledgements

This paper is supported in part by National Key R & D Plan of China (No. 2017YFB1002802), NSFC (No. 61572471), and Beijing Municipal Science & Technology Commission (No.Z171100000117017).

References

[Albinali *et al.*, 2010] Fahd Albinali, Stephen Intille, William Haskell, and Mary Rosenberger. Using wearable activity type detection to improve physical activity energy expenditure estimation. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 311–320. ACM, 2010.

[Andreu-Perez *et al.*, 2015] Javier Andreu-Perez, Daniel R Leff, Henry MD Ip, and Guang-Zhong Yang. From wearable sensors to smart implants—toward pervasive and personalized healthcare. *IEEE Transactions on Biomedical Engineering*, 62(12):2750–2762, 2015.

[Anguita *et al.*, 2012] Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, and Jorge L Reyes-Ortiz. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In *International workshop on ambient assisted living*, pages 216–223. Springer, 2012.

[Atkinson *et al.*, 2007] Hal H Atkinson, Caterina Rosano, Eleanor M Simonsick, Jeff D Williamson, Cralen Davis, Walter T Ambrosius, Stephen R Rapp, Matteo Cesari, Anne B Newman, Tamara B Harris, et al. Cognitive function, gait speed decline, and comorbidities: the health, aging and body composition study. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 62(8):844–850, 2007.

[Bonawitz *et al.*, 2017] Keith Bonawitz, Vladimir Ivanov, Ben Kreuter, Antonio Marcedone, H Brendan McMahan, Sarvar Patel, Daniel Ramage, Aaron Segal, and Karn Seth. Practical secure aggregation for privacy-preserving machine learning. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 1175–1191. ACM, 2017.

[Chen *et al.*, 2017] Yiqiang Chen, Xiaodong Yang, Biao Chen, Chunyan Miao, and Hanchao Yu. Pdassist: Objective and quantified symptom assessment of parkinson's disease via smartphone. In *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 939–945. IEEE, 2017.

[Chen *et al.*, 2018a] Fei Chen, Zhenhua Dong, Zhenguo Li, and Xiuqiang He. Federated meta-learning for recommendation. *arXiv preprint arXiv:1802.07876*, 2018.

[Chen *et al.*, 2018b] Yiqiang Chen, Chunyu Hu, Bin Hu, Lisha Hu, Han Yu, and Chunyan Miao. Inferring cognitive wellness from motor patterns. *IEEE Transactions on Knowledge and Data Engineering*, 30(12):2340–2353, 2018.

[Chen *et al.*, 2019] Yiqiang Chen, Jindong Wang, Meiyu Huang, and Han Yu. Cross-position activity recognition with stratified transfer learning. *Pervasive and Mobile Computing*, 57:1–13, 2019.

[Cheng *et al.*, 2019] Kewei Cheng, Tao Fan, Yilun Jin, Yang Liu, Tianjian Chen, and Qiang Yang. Secureboost: A lossless federated learning framework. *CoRR*, abs/1901.08755, 2019.

[Fernando *et al.*, 2013] Basura Fernando, Amaury Habrard, Marc Sebban, and Tinne Tuytelaars. Unsupervised visual domain adaptation using subspace alignment. In *Proceedings of the IEEE international conference on computer vision*, pages 2960–2967, 2013.

[Ganin and Lempitsky, 2015] Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. In *Interntional Conference on Machine Learning (ICML)*, 2015.

[Geyer *et al.*, 2017] Robin C Geyer, Tassilo Klein, and Moin Nabi. Differentially private federated learning: A client level perspective. *arXiv preprint arXiv:1712.07557*, 2017.

[Gong *et al.*, 2012] Boqing Gong, Yuan Shi, Fei Sha, and Kristen Grauman. Geodesic flow kernel for unsupervised domain adaptation. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2066–2073. IEEE, 2012.

[Hiremath *et al.*, 2014] Shivayogi Hiremath, Geng Yang, and Kunal Mankodiya. Wearable internet of things: Concept, architectural components and promises for person-centered healthcare. In *2014 4th International Conference on Wireless Mobile Communication and Healthcare-Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH)*, pages 304–307. IEEE, 2014.

[Huang *et al.*, 2007] Jiayuan Huang, Arthur Gretton, Karsten Borgwardt, Bernhard Schölkopf, and Alex J Smola. Correcting sample selection bias by unlabeled data. In *Advances in neural information processing systems*, pages 601–608, 2007.

[Huang *et al.*, 2012] Pipei Huang, Gang Wang, and Shiyin Qin. Boosting for transfer learning from multiple data sources. *Pattern Recognition Letters*, 33(5):568–579, 2012.

[Inkster, 2018] Nigel Inkster. *China's Cyber Power*. Routledge, 2018.

[Konečnỳ *et al.*, 2016] Jakub Konečnỳ, H Brendan McMahan, Daniel Ramage, and Peter Richtárik. Federated optimization: Distributed machine learning for on-device intelligence. *arXiv preprint arXiv:1610.02527*, 2016.

[Lei *et al.*, 2012] Jinna Lei, Xiaofeng Ren, and Dieter Fox. Fine-grained kitchen activity recognition using rgb-d. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 208–211. ACM, 2012.

[Menshawy *et al.*, 2015] Mohamed EL Menshawy, Abdelghani Benharref, and Mohamed Serhani. An automatic mobile-health based approach for eeg epileptic seizures detection. *Expert Systems with Applications*, 42(20):7157–7174, 2015.

[Michalak *et al.*, 2009] Johannes Michalak, Nikolaus F Troje, Julia Fischer, Patrick Vollmar, Thomas Heidenreich, and Dietmar Schulte. Embodiment of sadness and depression—gait patterns associated with dysphoric mood. *Psychosomatic medicine*, 71(5):580–587, 2009.

[Pan and Yang, 2010] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2010.

[Pan *et al.*, 2011] Sinno Jialin Pan, Ivor W Tsang, James T Kwok, and Qiang Yang. Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks*, 22(2):199–210, 2011.

[Rashidi and Cook, 2010] Parisa Rashidi and Diane J Cook. Mining sensor streams for discovering human activity patterns over time. In *2010 IEEE International Conference on Data Mining*, pages 431–440. IEEE, 2010.

[Rebuffi *et al.*, 2017] Sylvestre-Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl, and Christoph H. Lampert. icarl: Incremental classifier and representation learning. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[Rivest *et al.*, 1978] Ronald L Rivest, Len Adleman, Michael L Dertouzos, et al. On data banks and privacy homomorphisms. *Foundations of secure computation*, 4(11):169–180, 1978.

[Rozantsev *et al.*, 2018] Artem Rozantsev, Mathieu Salzmann, and Pascal Fua. Beyond sharing weights for deep domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41:801–814, 2018.

[Rozantsev *et al.*, 2019] Artem Rozantsev, Mathieu Salzmann, and Pascal Fua. Beyond sharing weights for deep domain adaptation. *IEEE transactions on pattern analysis and machine intelligence*, 41(4):801–814, 2019.

[Shokri and Shmatikov, 2015] Reza Shokri and Vitaly Shmatikov. Privacy-preserving deep learning. In *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*, pages 1310–1321. ACM, 2015.

[Smith *et al.*, 2017] Virginia Smith, Chao-Kai Chiang, Maziar Sanjabi, and Ameet S Talwalkar. Federated multi-task learning. In *Advances in Neural Information Processing Systems*, pages 4424–4434, 2017.

[Sun *et al.*, 2016] Baochen Sun, Jiashi Feng, and Kate Saenko. Return of frustratingly easy domain adaptation. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

[Tzeng *et al.*, 2014] Eric Tzeng, Judy Hoffman, Ning Zhang, Kate Saenko, and Trevor Darrell. Deep domain confusion: Maximizing for domain invariance. *arXiv preprint arXiv:1412.3474*, 2014.

[Voigt and Von dem Bussche, 2017] Paul Voigt and Axel Von dem Bussche. The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, 2017.

[Wang *et al.*, 2014] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 3–14. ACM, 2014.

[Wang *et al.*, 2017a] Jindong Wang, Yiqiang Chen, Shuji Hao, Wenjie Feng, and Zhiqi Shen. Balanced distribution adaptation for transfer learning. In *2017 IEEE International Conference on Data Mining (ICDM)*, pages 1129–1134. IEEE, 2017.

[Wang *et al.*, 2017b] Yuxi Wang, Kaishun Wu, and Lionel M Ni. Wifall: Device-free fall detection by wireless networks. *IEEE Transactions on Mobile Computing*, 16(2):581–594, 2017.

[Wang *et al.*, 2018a] Jindong Wang, Yiqiang Chen, Lisha Hu, Xiaohui Peng, and S Yu Philip. Stratified transfer learning for cross-domain activity recognition. In *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 1–10. IEEE, 2018.

[Wang *et al.*, 2018b] Jindong Wang, Wenjie Feng, Yiqiang Chen, Han Yu, Meiyu Huang, and Philip S Yu. Visual domain adaptation with manifold embedded distribution alignment. In *2018 ACM Multimedia Conference on Multimedia Conference*, pages 402–410. ACM, 2018.

[Wang *et al.*, 2019a] Jindong Wang, Yiqiang Chen, Shuji Hao, Xiaohui Peng, and Lisha Hu. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 119:3–11, 2019.

[Wang *et al.*, 2019b] Jindong Wang, Yiqiang Chen, Han Yu, Meiyu Huang, and Qiang Yang. Easy transfer learning by exploiting intra-domain structures. In *IEEE International Conference on Multimedia and Expo (ICME)*, 2019.

[Ward *et al.*, 2016] Jamie A Ward, Gerald Pirkl, Peter Hevesi, and Paul Lukowicz. Towards recognising collaborative activities using multiple on-body sensors. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 221–224. ACM, 2016.

[Yang *et al.*, 2018] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. Federated learning. *Communications of the CCF*, 11:49–55, 2018.

[Yang *et al.*, 2019] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(2):12, 2019.

[Yosinski *et al.*, 2014] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? In *Advances in neural information processing systems*, pages 3320–3328, 2014.

[Zheng *et al.*, 2018] Zibin Zheng, Shaoan Xie, Hong-Ning Dai, Xiangping Chen, and Huaimin Wang. Blockchain challenges and opportunities: a survey. *International Journal of Web and Grid Services*, 14(4):352–375, 2018.