Deep Learning Enhanced Human Activity Recognition for Internet of Healthcare Things

Xiaokang Zhou, *Member, IEEE,* Wei Liang, *Member, IEEE,* Kevin I-Kai Wang, *Member, IEEE,* Hao Wang, *Member, IEEE,* Laurence T. Yang, *Fellow, IEEE*, and Qun Jin, *Senior Member*, *IEEE*

***Abstract*—Along with the advancement of several emerging computing paradigms and technologies, such as cloud computing, mobile computing, artificial intelligence, and big data, Internet of Things (IoT) technologies have been applied in a variety of fields. In particular, the Internet of Healthcare Things (IoHT) is becoming increasingly important in Human Activity Recognition (HAR), due to the rapid development of wearable and mobile devices. In this study, we focus on the deep learning enhanced HAR in IoHT environments. A semi-supervised deep learning framework is designed and built for more accurate HAR, which efficiently use and analyze the weakly labeled sensor data to train the classifier learning model. To better solve the problem of inadequately labeled sample, an intelligent auto-labeling scheme based on Deep Q-Network (DQN) is developed with a newly designed distance-based reward rule, which can improve the learning efficiency in IoT environments. A multi-sensor based data fusion mechanism is then developed to seamlessly integrate the on-body sensor data, context sensor data, and personal profile data together, and a Long Short Term Memory (LSTM)-based classification method is proposed to identify fine-grained patterns according to the high-level features contextually extracted from sequential motion data. Finally, experiments and evaluations are conducted to demonstrate the usefulness and effectiveness of the proposed method using real world data.**

***Index Terms*—Human Activity Recognition, Deep Learning, Reinforcement Learning, Internet of Things, Weakly Labeled Data, Smart Healthcare**

I. Introduction

The rapid development of wearable and mobile devices has facilitated the application of Internet of Things (IoT) technologies in healthcare field. Monitoring real time human activities, especially Activities of Daily Living (ADL) of elder people, is an essential issue in smart healthcare, which can evidently enhance the medical rehabilitation and elderly care using wearable and mobile sensors. Accordingly, Human Activity Recognition (HAR) in ubiquitous computing environments, has become a hotly discussed topic in better understanding people's daily behaviors and interactions with their living environments, which is studied extensively for the so-called Internet of Healthcare Things (IoHT) [1]. Different kinds of wearable devices (e.g., body-worn inertial sensors, and smartphone) can be placed on different on-body locations (e.g., head, chest, upper arm, forearm, shin, etc.) to collect and transfer real time posture data (e.g., accelerometer and gyroscope) using wireless sensor networks [2]. These sensor technologies provide us opportunities to improve the robustness of multi-modal data sensing and fusion in HAR, which may support the development of human-centric applications and services in cyber-physical-social systems based on the enrichment of sensed information from real-time big data environments [3, 4].

Since the motion data based on people's physical activities can be easily obtained via built-in inertial sensors of wearable and mobile devices, activity recognition and classification using smart devices are widely used to monitor, analyze, and understand one person's status in different scenes across a variety of applications and systems [5, 6]. However, there are still several challenges in HAR via sensor data. For example, the effectiveness of traditional machine learning based methods mainly rely on the availability of training data, which means feature extraction mostly depends on a well-designed dataset with adequate labeling [7, 8]. However, the ADL data is continuously generated by the built-in sensors of smart devices, regardless of whether people are performing sensible actions, which results in a labor intensive process of annotating and recording well-labeled data. This kind of wearable sensor data, namely, massive amount of unlabeled data combined with small portion of labeled data, can be referred to as weakly labeled motion data [9, 10]. It is necessary to build a new semi-supervised learning or weakly supervised learning framework, to efficiently deal with such kind of situation.

Actually, it is obvious that activity recognition in real world is depended on the subjects and where the wearable device is worn, and signals obtained by on-body sensors is up to the

X. Zhou is with the Faculty of Data Science, Shiga University, Hikone, and RIKEN Center for Advanced Intelligence Project, Tokyo, Japan (e-mail: zhou@biwako.shiga-u.ac.jp).

W. Liang (corresponding author) is with the Key Laboratory of Hunan Province for New Retail Virtual Reality Technology, Hunan University of Technology and Business, and Business School, Central South University, Changsha, China (e-mail: weiliang@csu.edu.cn).

K. Wang is with the Department of Electrical, Computer and Software Engineering, The University of Auckland, Auckland, New Zealand (e-mail: kev-in.wang@auckland.ac.nz).

H. Wang is with the Department of Computer Science, Norwegian University of Science & Technology, Norway (e-mail: hawa@ntnu.no).

Laurence T. Yang is with the Department of Computer Science, St. Francis Xavier University, Antigonish, Canada (e-mail: ltyang@ieee.org)

Q. Jin is with the Faculty of Human Sciences, Waseda University, Tokorozawa, Japan (e-mail: jin@waseda.jp).

positions of them. Different positions of on-body sensors can generate different motion patterns even for the same activity performed by one person. Existing works are able to recognize coarse-grained behaviors from repetitive movements (e.g., walking, running), static actions (e.g., sitting, standing), or simple transitional activities (e.g., stand-sit, sit-lie) with relatively high accuracy, but face challenges to identify some complex patterns with single sensor data [11, 12]. Sensor data collected from multiple wearable devices, including accelerometer, GPS, light, video, etc., needs to be synchronized and integrated together through a unified data fusion strategy, to capture more complex human activity patterns from a multi-modal and multi-positional view. In addition, with the same on-body sensors, different subjects may produce different motion patterns for the same type of activity. For instance, the falling pattern of an elder person and a child can be entirely different. Moreover, traditional approaches on feature extraction for HAR may only figure out low-level features, which could be enough to recognize basic physical or postural activities. But without considerations of some context-aware or location-aware issues, it would be a challenge to detect more meaningful actions with semantic information, such as jogging, cooking, etc. Therefore, it is essential to find an efficient way to extract the high-level features within a certain context, which may facilitate pattern recognitions in terms of the fine-grained human actions, gestures, and expressions.

In this study, we focus on the deep learning enhanced activity recognition, in order to detect the fine-grained motion patterns through better utilization of weakly labeled sensor data. Considering the multiple sensor data obtained by various wearable devices, a semi-supervised learning framework is introduced to incorporate massive amount of unlabeled data with small portion of labeled data to enhance the accuracy of HAR in IoHT environments. Specifically, comparing with the existed related researches, contributions of our study can be concluded as follows.

i) A semi-supervised deep learning framework is designed and constructed with an auto-labeling module and a Long Short Term Memory (LSTM)-based classification module, which can efficiently utilize the massive amount of weakly labeled data to train the classifier, and improve the accuracy of HAR in IoHT environments.
ii) An intelligent auto-labeling scheme based on Deep Q-Network (DQN) is developed with a novel distance-based reward rule, which can better solve the problem of inadequately labeled sample, and improve the learning efficiency.
iii) An LSTM-based classifier is built with a multi-sensor based data fusion mechanism, which can be used to deal with the sequential motion data and detect fine-grained patterns according to the extracted high-level features.

The rest of this article is organized as follows. Section II presents an overview of related works. The framework of semi-supervised deep learning model is presented in Section III. Algorithm and mechanism for auto-labeling from weakly labeled data and fine-grained activity recognition are discussed in Section IV. In Section V, we discuss the experiment and evaluation results using two real world datasets. In Section VI, we conclude this research and give a promising perspective on future research.

II. Related Work

The study of HAR is an important research direction in the field of IoHT. Research works in this direction are applied in various practical applications such as healthcare, gym physical activity recognition, fall detection, etc. Several issues relating to this topic are discussed in this section. Foremost, studies on HAR, issues of analysis based on multiple motion data, and researches on semi-supervised learning for HAR, are addressed respectively.

A. Studies on HAR

HAR can be viewed as one kind of artificial intelligent technology, which analyzes and recognizes human activities and behavior patterns automatically through a series of observations of wearable device data. Generally, Turaga et al. [7] divided HAR into three levels as: the movement recognition, action recognition and activity recognition, which were referred to as the low-level vision [13], middle-level vision [14], and high-level vision respectively [15]. Different machine learning and deep learning based schemes were explored to handle issues in HAR and achieved effective performance [16].

Recently, studies focusing on HAR could be classified into two important categories as ambient sensor based and wearable sensor based approaches. Ambient sensor based approaches usually deployed surveillance camera, sound, temperature and other indoor sensors to capture environment-related context signals, and recognize people's daily activities in the fixed space (e.g., smart home [17], recovery center [18]). The requirement of fixed environments made them not applicable for analyzing normal outdoor activities. On the other hand, wearable sensor based approach utilized wearable devices or smartphones to monitor and obtain on-body physiological signals with accelerometer, magnetometer and gyroscope sensors [19]. For example, Lee and Mase [20] used the acceleration and angular velocity sensor data measured by inexpensive wearable devices, to determine the users' location information and identify their sitting, standing and walking behaviors. Mantyjarvi et al. [21] utilized the independent component analysis and principal component analysis schemes to recognize one person's walking posture based on the acceleration data collected from the buttocks. These results demonstrated that activity recognition techniques based on wearable sensors could work effectively for the low-level vision, but failed to handle high-level recognition tasks for complex activity recognition [22].

B. Analysis Based on Multiple Motion Data

Typically, lots of studies focused on human activity recognition with the accelerometer sensor data. Bao and Intille [23] developed algorithms to detect physical activities from subjects' daily tasks using multiple accelerometers. Kern et al. [24] placed a three-axis accelerometer on different parts of the user's body, where each accelerometer could reflect different directions and motions of relevant parts. However, these approaches were not easy to achieve a higher performance because of the constraint of single data source without enough

context information. Lukowicz et al. [25] used gyroscopes and accelerometers to measure the acceleration and angular velocity. But the recognition accuracy was affected by the limited experimental conditions, especially by the noisy data and sensor variations. Patterson et. al. [26] presented a Bayesian model with traveler's moving data drawn from the GPS sensor stream, to learn travelers' transportation mode and their most likely route in an unsupervised manner. They demonstrated that the recognition accuracy could be improved by adding external context knowledge about bus routes and bus stops. Liao et al. [27] introduced a hierarchical Markov model to infer one user's daily movements by integrating raw GPS sensor measurements with high-level information (e.g., user's destination, mode of transportation), which was applied to help people use the public transportation more safely. In particular, several existing HAR methods explored hybrid classifiers to improve recognition performance [28], which were usually trained and combined multiple classifiers from different sensor data according to their corresponding feature patterns, in IoT and network computing environments [29, 30]. However, these methods had limited considerations of unlabeled data.

C. Semi-Supervised Learning for HAR

Supervised learning model was widely implemented for HAR, in which the collected labeled data was used to train the model. However, most data collected by wearable devices was unlabeled, which led to a tedious and costly work of annotating those unlabeled data. This has become a difficult issue for HAR in IoHT environment [31].

Semi-supervised learning is a learning approach that combines supervised learning with unsupervised learning techniques, which is studied for human activity recognition extensively. It can incorporate the large portion of unlabeled data with labeled data to solve the problem of inadequate annotation for human activities. Motivated by the classification problem of web pages, Blum et al. [32] proposed a co-training framework to leverage the unlabeled data to enlarge the training set, and improve the performance of the recognition algorithm. Zhou et al. [33] proposed a so-called disagreement-based semi-supervised learning paradigm, in which multiple classifiers were trained from real-world tasks, and disagreements among the classifiers were exploited to guide the semi-supervised learning process. Zhu et al. [9] proposed a semi-supervised deep learning approach, in which the temporal ensemble of deep LSTM was developed to recognize human activities with labeled and unlabeled smartphone inertial sensor data. The unsupervised losses were leveraged and combined together with the supervised losses for the accurate HAR. To investigate unobtrusive and context-aware activity recognition using on-body wearable sensors, Stikic et al. [34] proposed an annotation strategy which leveraged sparsely labeled data together with more easily obtainable unlabeled data.

III. Framework of Semi-Supervised Deep Learning

In this section, after introducing the formal description of HAR problem in IoHT environments, we present the design of a semi-supervised deep learning framework with its core function modules.

A. Problem Definition

HAR for daily living in the IoHT environment can be viewed as a time-dependent task since the data is generated continuously via wearable devices. The main challenge in designing the classifier is how to reasonably segment this time stream data, which is also an essential step to extract features for each activity. Recurrent Neural Network (RNN) [35] is one kind of classic supervised learning model, which can utilize the internal state memory to record the temporal information between layers of a neural network. This allows to process the unsegmented and temporal sequence data for applicable tasks in real world, such as nature language process, handwriting recognition, speech recognition, and HAR, etc. In particular, LSTM model is a special kind of recurrent network that incorporates some gates and memory cells into the network, in order to capture the long-term dependences of the sequence data. In this study, we utilize the LSTM model in the constructed deep learning framework, to detect the temporal features of segmented motion activities.

As we discussed earlier, labeling for the accelerometer and gyroscope data generated by wearable devices is a costly and labor-intensive task, which indicates the importance of using weakly labeled data to design a practical classifier for HAR. However, it may result in unfavorable results if we simply apply the traditional LSTM model to deal with such kind of weakly labeled data in HAR, because it contains massive unlabeled data mixed with only a small portion of labeled data. Therefore, a semi-supervised deep learning framework is designed to handle this situation.

Given the HAR problem in IoHT environments with a time series of unlabeled and labeled data uniformly indicated as X_t with the timestamp *t*, obtained by different kinds of on-body wearable devices, the definition of labeled data can be expressed as follows.

$$LB(X\_lb_t, Y\_lb_t, S, D) \tag{1}$$

Where $X\_lb_t = \{x\_lb_t^n | n \in \{1,2,\dots,N\}\}$ denotes the labeled dataset, $X\_lb_t \subset X_t$, and $Y\_lb_t = \{y\_lb_t^n | n \in \{1,2,\dots,N\}\}$ denotes their corresponding label set, N is the number of the labeled data. $S = \{s_1, s_2, \dots, s_h\}$ denotes the subject set, and $D = \{d_1, d_2, \dots, d_f\}$ stands for the set of wearable devices used by the subjects, in which each device d_k stands for a typical on-body devices used by a subject s_i.

Likewise, the definitions of unlabeled data can be expressed as follows.

$$ULB(X\_ulb_t, S, D) \tag{2}$$

where $X\_ulb_t = \{x\_ulb_t^m | m \in \{1,2,\dots,M\}\}$ denotes the unlabeled dataset, $X\_ulb_t \subset X_t$. Note that we assume $N \ll M$ in this study.

Thus, given a set of test data $X_t' \subset X_t$ in a certain temporal sequence, the problem studied here is to detect the corresponding activities implicated in X_t', i.e., recognize the activity annotation Y_t' for X_t' based on the fine-grained feature extraction and classification in a context-aware way. Furthermore, to address the detailed attributes of each subject, a four-dimension tuple is defined to describe their profile as follows.

$$Pro_i = (gd_i, ag_i, ht_i, wt_i) \tag{3}$$

where gd_i, ag_i, ht_i, and wt_i denote a specific subject's gender, age, height, and weight respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2020.2985082, IEEE Internet of Things Journal

B. Framework Overview

Following definitions we introduced above, a semi-supervised deep learning framework is designed and constructed to build the fine-grained classifier for HAR in IoHT environments, which is shown in Fig. 1.

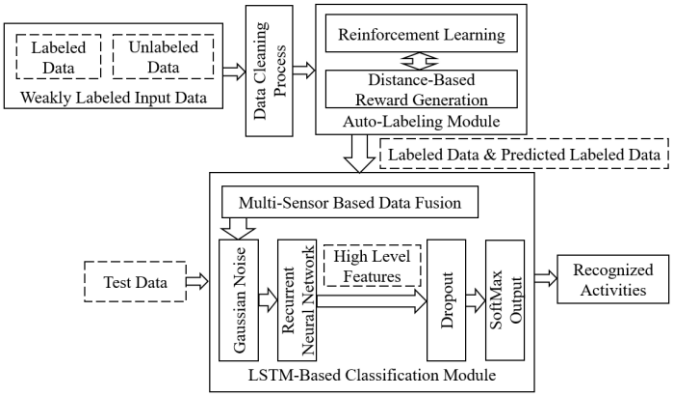


Fig. 1. Framework of Semi-supervised deep learning for HAR with weakly labeled input data

As shown in Fig. 1, the whole framework is mainly composed of two basic modules, namely, the auto-labeling module and the LSTM-based classification module. More precisely, multiple types of data, including the on-body sensor data, context sensor data, and personal profile data, are selected as the input to the auto labeling and classification module. The data cleaning process is performed firstly to handle the incomplete, incorrect, or irrelevant parts of the input data, after which the pre-processed data is sent to the auto-labeling module.

In the auto-labeling module, a reinforcement learning based auto-segmentation and labeling process is applied before sending the weakly labeled input data to the LSTM neural network. Given X_t, the goal of the auto labeling module is to convert all the samples in ULB to LB. Specifically, assuming the state of an input sample data x_t at timestamp t is st_t, an agent is designed to conduct a possible action act_t to assign a label y_t for x_t. The core idea of the auto-labeling is to assign rewards for actions conducted by the agent. In this study, a distance-based reward rule is designed to enable the agent to choose a correct action with more confidence. Each time when the agent succeeded in predicting a correct label, it will receive a positive reward $1 * \gamma$. In contrast, if the agent fails to predict the correct label, it will receive a penalty of $-1 * \gamma$, where γ is the distance-based weight parameter.

In the classification module, a multi-sensor based data fusion technique is developed to integrate multiple on-body sensor data together to handle complex activity recognition tasks. Considering the case that detecting falling patterns of an elder and a child based on their motions respectively, fusion of multiple sensor data captured from different body positions may provide enriched motion information to distinguish this similar pattern recognition. Furthermore, a K layer LSTM based deep learning network is adopted to extract the high-level features based on the fused multi-sensor data. Given a set of weakly labeled input data X_t, the core components of LSTM can be expressed as follows.

$$
\begin{aligned}
f_t &= \sigma(W_f \cdot [h_{t-1}, X_t] + b_f) \\
i_t &= \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \\
\tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, X_t] + b_C) \\
C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t \\
o_t &= \sigma(W_o \cdot [h_{t-1}, X_t] + b_o) \\
h_t &= o_t \cdot \tanh(C_t)
\end{aligned} \tag{4}
$$

where $f_t,\ i_t, o_t,\ \tilde{C}_t, C_t, h_t$ stand for the forget gate, input gate, output gate, two memory cells for state persistence, and the hidden layer respectively. $\sigma(*)$ is a sigmoid activation function to introduce nonlinear variations to the network. $W_f,\ W_i, W_C,\ W_o$ are the weights for each component, and $b_f,\ b_i, b_C,\ b_o$ are the biases.

In addition, to avoid the overfitting issue and enhance the generalization for the deep learning network, the Gaussian noise is included in the classification module, and the dropout component with probabilities p_{drop} is applied in the LSTM network. The SoftMax function is utilized to handle the multi-classification situation.

IV. Mechanism for Semi-Supervised Human Activity Recognition

In this section, we discuss the auto-labeling propagation process incorporating the reinforcement learning model to resolve the HAR problem with weakly labeled data. A DQN-based auto-labeling scheme is introduced, a multi-sensor based data fusion mechanism, and a LSTM-based fine-grained classifier for HAR are developed.

A. Reinforcement Learning Based Auto-Labeling Scheme

Considering the challenge of HAR based on weakly labeled data in IoHT environments, we develop an intelligent auto-labeling scheme based on reinforcement learning model to find an optimal strategy that can maximize the reward during the labeling process. DQN, which is a combination of Q-learning and deep neural network, is capable of solving the complex problem with uncertainty.

Basically, the auto-labeling scheme based on reinforcement learning, can be described according to an agent and its triple (st_t, act_t, rwd_t) at timestamp t, in which st_t denotes the current agent's state at t, act_t denotes a current action adopted by the learning model, and rwd_t denotes the reward determined by st_t and act_t. Additionally, to describe the auto-labeling process, a reward function $R(st_t, act_t, \gamma)$ is defined to evaluate the mapping from state to action, which can be expressed as Eq. (5).

$$R(st_t, act_t, \gamma) = E[rwd_{t+1} + \gamma\, Q^{\pi}(st_t, act_t, \theta) \mid st_t, act_t] \tag{5}$$

The Q-function $Q^{\pi}(st_t, act_t, \theta)$ is introduced to evaluate the cumulative reward expectation gained by the agent according to st_t and act_t, and can be described as Eq. (6).

$$Q^{\pi}(st_t, act_t, \theta) = E_{\pi}\{\textstyle\sum_{k=0}^{\infty} \gamma^k\, rwd_{t+k} \mid st_t, act_t\} \tag{6}$$

where θ is the DQN parameter, $\gamma \in [0,1]$ is the discount parameter.

The key issue of auto-labeling is to design a reasonable reward rule, i.e., the reward function defined in Eq. (5), to propagate labels for massive unlabeled data with high accuracy. Specifically, the point is how to design the discount parameter γ for the reward function. Normally, the data with identical label usually may have similar attributes, we thus try to deduce

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2020.2985082, IEEE Internet of Things Journal

γ for each propagation action by evaluating the distance of an unlabeled input data against the nearest labeled neighbor, which means we cluster all the data by the unsupervised clustering method firstly, and then measure the distance of each unlabeled data to the nearest cluster with labeled data. Concretely, given a set of input data $X_t = X\_lb_t \cup X\_ulb_t$ at timestamp t, the dataset can be firstly clustered into k clusters as $C = \{c_1, c_2, \dots, c_k\}$. Then, for each unlabeled data $x_t^m \in X\_ulb_t, m \in \{1,2, \dots, M\}$, assuming its nearest cluster containing labeled data is c_k, thus the weight of its reward/penalty is calculated by Euclidean Distance between x_t^u to the corresponding cluster center of c_k. The detailed calculation of the propagation entropy of the labeling can be expressed as follows.

$$\gamma_{lk} = -\ln(dis(x_l, c_k)) \tag{7}$$

Based on these, if the propagation entropy value is more than a given threshold, x_t^m will be assigned with the label from this cluster c_k. Finally, the goal of the auto-labeling is to find an optimal action act_t^* at each step by estimating the state st_t at t. The concrete auto-labeling scheme is illustrated in Fig 2.

Input: The weakly labeled data $LB \cup ULB$;
The clustering set $C = \{c_1, c_2, \dots, c_k\}$
Output: A set of labeling actions Act

Initialize action-value function Q with random weight θ
Initialize target action-value function $\hat{Q}$ with θ
Initialize parameters: *EP, T*
for episode = 1 to *EP* **do**
  Initialize sequence $St = \{st_1\}$, and preprocess $\Phi_1 = st_1$
  for t = 1 to T **do**
    Select a random action act_t, $Act = Act \cup act_t$
    Execute action act_t and observe the responding reward $rwd_t(st_t, act_t, \gamma)$
    Set $st_{t+1} = st_t, act_t$ and preprocess $\Phi_{t+1} = \Phi(st_{t+1})$
    Sample random minibatch of transitions $(\Phi_j, act_j, rwd_j, \Phi_{j+1})$
    Set $y_j =$
$$\begin{cases} rwd_j & if\ episode\ terminates\ at\ step\ j+1 \\ R(\Phi_j, act, \gamma) \text{ by Eq.(5)} & otherwise \end{cases}$$
    Perform a gradient descent step on $(y_j - Q(\Phi_j, act_j; \theta))^2$ with respect to weight θ
    Every 100 steps reset $\hat{Q} = Q$
  end for
end for
Return Act

Fig. 2. Scheme for auto-labeling based on DQN

B. Multi-Sensor Based Data Fusion Mechanism

The on-body sensor data, context sensor data, and personal profile data are integrated together to figure out different motion patterns in IoHT environments.

We utilize on-body devices in six different positions to capture subjects' accelerometer data simultaneously. The on-body positions of the sensors are listed in Table 1.

TABLE I
DESCRIPTION OF ON-BODY DEVICES

Devices	On-body position	Devices	On-body position
d_1	head	d_4	waist
d_2	chest	d_5	thigh
d_3	upper arm	d_6	shin

Given a subject s_i, we build the tensor for on-body accelerometer sensor data, and integrate the above six sensors as in Eq. (8).

$$data_i^{acc}(data_{i,d_1}, data_{i,d_2}, \dots, data_{i,d_6}) \tag{8}$$

Three kinds of modalities of context sensor data, i.e., GPS, orientation, light, are adopted to enable the context-aware HAR in the fine-grained classifier. Given a subject s_i, we build the tensor for context sensor data as in Eq. (9).

$$data_i^{cot}(gps_i, ort_i, lit_i) \tag{9}$$

Accordingly, the final tensor for data fusion representation is integrated and constructed as follows.

$$z_i(s_i, data_i^{acc}, data_i^{cot}) \tag{10}$$

A simple data sample, including the personal profile $Pro_i = (gd_i, ag_i, ht_i, wt_i)$, accelerometer data $data_i^{acc}$, and context data $data_i^{cot}$, obtained by devices $\{d_k\}$ from subjects $\{s_i\}$, can be exemplified in Table II with two subjects s_1, s_2 and three on-body devices d_1, d_2, d_3 considering the location context.

TABLE II
MULTI-SENSOR DATA SAMPLE

Subjects	Profiles	Devices	Accelerometer	GPS
s_1	F, 28, 165, 65	d_1	49.48, 49.48, …	42.2, -71.3
s_1	F, 28, 165, 65	d_2	8.47,8.45, …	41.4, -72.4
s_1	F, 28, 165, 65	d_3	-0.92, -1.04, …	39.8, -65.7
s_2	M, 45, 174, 76	d_1	42.48, 42.75, …	86.1, 4.3
s_2	M, 45, 174, 76	d_2	9.62, 9.56, …	80.7, 4.1
s_2	M, 45, 174, 76	d_3	-1.04, -1.05, …	78.5, 3.9

C. LSTM-Based Fine-Grained Activity Recognition in IoHT

We apply the LSTM recurrent network to capture temporal features hidden in the fused multi-sensor data. As illustrated in Fig. 1, we send a sequence of fused multi-sensor data constructed as Eq. (10) with their corresponding labels into the LSTM-based deep neural network, in order to train the fine-grained classification model. The concrete training process is illustrated in Fig. 3.

Input: A sequence of fused input data $Z = \{z_1, z_2, \dots, z_n\}$
A set of labels Y corresponding to Z
Output: A trained human activity classification model

Initialize parameter $Iter, Batch, Thresh_{loss}$
for iteration = 1 to $Iter$ **do**
  for each minibatch $\{z_k\}_{k=1}^{n/Batch}$ **do**
    Filter input data $g_k = \mathrm{G}(z_k)$ with Gaussian Noise layer
    Compute high level feature $hf_k = \mathrm{LSTM}(g_k)$ with LSTM recurrent network
    Conduct the variation computation with dropout 0.5: $dp_k = \mathrm{dropout}_{0.5}(g_k)$
    Predict the activity label $\tilde{y}_k = \mathrm{SoftMax}(dp_k)$
    Compute the cost function about the loss by $\tilde{y}_k$ and $y_k \in Y$
    if *loss* < $Thresh_{loss}$: **break**
    Run the Back-Propagation process to update the parameters for the model
  end for
end for

Fig. 3. Fine-grained classification training based on LSTM

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2020.2985082, IEEE Internet of Things Journal

As shown in Fig. 3, with a sequence of fused input data $Z = \{z_1, z_2, \dots, z_n\}$ and a set of corresponding labels (Note that some of these labels are propagated through the proposed auto-labeling scheme), we start the training process upon the pre-setting parameters, i.e., the max iterations M, batchsize $Batch$, and the threshold for cost function $Thresh_{loss}$. Each batch of the fused input data runs through the Gaussian noise, LSTM recurrent neural network, Dropout and SoftMax output layers in the forward-propagation process. Then, the training loss is calculated by the cost function and applied in the back-propagation process to update the model. When the loss is lower than the threshold, the model is prepared and ready for HAR in IoHT environments.

V. Experiment And Analysis

To verify the effectiveness of our method to the fine-grained human activity recognition in IoHT, comparison experiments are conducted with multiple sensor data collected in real world. All experiments have run on a server of Intel i7-4790 @3.6GHz CPU, 32GB RAM, NVidia GeForce GTX 970 GPU, Linux, Python 3.5, TensorFlow r2.0.

A. Data Set

To investigate the proposed method on HAR in IoHT environments, two free datasets [8, 36], which were collected using smartphones and on-body wearable devices (see Table I), were downloaded for the comparison evaluation. In the IoHT environment, the on-body wearable devices were communicated and synchronized with the network provider.

The first dataset was a pure acceleration data, which included 11,771 activities performed by 30 subjects of ages ranging from 18 to 60 [8]. Specifically, this dataset covered 9 types of activities in daily living and 8 types of falls. Another dataset was collected by a set of on-body wearable devices, which also included several context information (e.g., GPS, magnetic field, sound level, and light) [36]. This dataset contained the acceleration data covering 8 types of activities from fifteen subjects (seven females and eight males, age 31.9 ± 12.4, height 173.1 ± 6.9, weight 74.1 ± 13.8). For each activity, the on-body positions: chest, forearm, head, shin, thigh, upper arm, and waist were simultaneously recorded.

B. Experiment Design and Evaluation Metrics

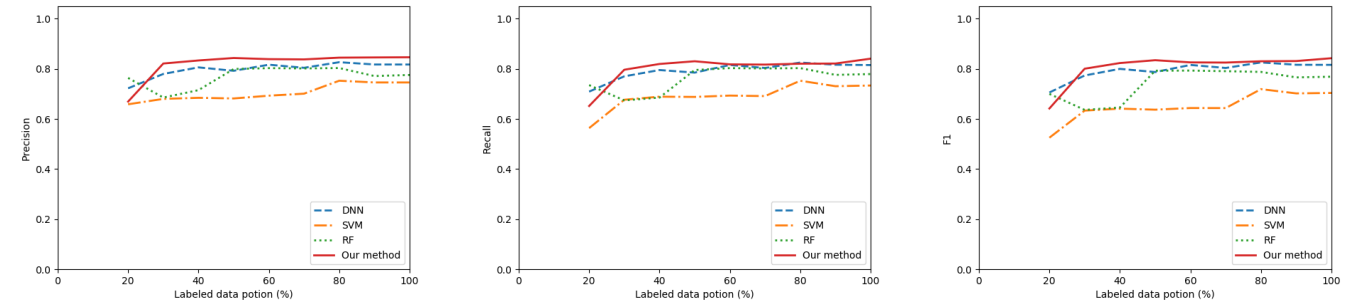
The whole dataset was separated into three parts: 60% as training subset, 20% as validation subset, and 20% as testing subset. Both the accelerometer and the corresponding context information in the train set were utilized to construct the Semi-Supervised learning model, and train the parameters in DQN and LSTM. To avoid the overfitting problem, the validation subset was applied to adjust the structure and parameters of the model.

We used the metrics: Accuracy, Precision-Recall, F1-score to evaluate the performance of the proposed method. We set a set of different portions of unlabeled data ranging from 20% to 100%. The DNN, RF, and SVM-based classifier are chosen for performance comparison for HAR in IoHT environments.

In addition, we used the Q-Learning and full connection deep neural network in DQN, the Gradient Descent Optimizer and SoftMax regression were used after the full connection layer to generate the classification results. A dropout scheme (dropout rate =0.4) was applied to avoid the overfitting problem. The learning rate was set to 0.05 and batch size was 25. We reset the Q function value by every 100 steps within an episode.

C. Effectiveness Evaluation for Auto-Labeling

We evaluated the effectiveness of the proposed auto-labeling method based on the weakly labeled data in IoHT environments.



(a) Precision (b) Recall (c) F1-score

Fig. 4. Efficiency evaluation on auto-labeling tasks.

We simulated the scenario that lacking of well labeled activities in daily living by moving out some of the labels from the dataset. Evaluations were conducted with the portions of unlabeled data ranging from 20% to 100%, to demonstrate the performance of our proposed method with the other three baseline methods under different degrees of weakly labeled data. Comparison results according to Precision, Recall, and F1-score are shown in Fig. 4 (a)-(c) respectively.

As shown in Fig. 4, generally, all the methods perform well in the fully supervised situation (100% portion of labeled data). However, results of all the methods becomes worse along with the increasing portion of unlabeled data. Since the proposed semi-supervised framework is benefited by our auto-labeling scheme, it achieves an average F1-score of 0.79 under different portion settings of unlabeled data.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2020.2985082, IEEE Internet of Things Journal

D. *Performance Evaluation for HAR*

We investigated eight types of activities in daily living, to evaluate the different performances for activity recognition, which are listed in Table III.

TABLE III
TYPICAL CATEGORIES OF ACTIVITIES TESTED IN THE EXPERIMENT

ID of activities	Class	Activity description
A_1	class 0	**Climbing down**
A_2	class 1	**Climbing up**
A_3	class 2	**Jumping**
A_4	class 3	**Sitting**
A_5	class 4	**Standing**
A_6	class 5	**Lying**
A_7	class 6	**Walking**
A_8	class 7	**Jogging**

First, we investigated how the proposed method performed with different kinds of activities. Fig. 5 demonstrates the recognition performance of the proposed method on different categories of activities (classes). We utilize the ROC curve to demonstrate the performance for different activities.

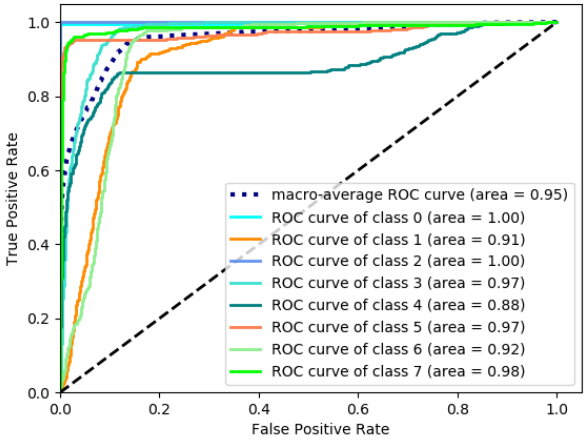


Fig. 5. Comparison result for all activities

The eight solid lines with different colors represent the Area Under Curve (AUC) values for activities listed in Table III. The dot line summarizes the macro-average performance of our proposed method among all the activities. The average AUC value is 0.95. In particular, recognition based on Climbing down (A_1) obtain the best performance, while recognition based on Standing (A_5) has the worst performance. This is mainly because the Climbing down activity may result in some distinctive features, while the Standing activity has fewer number of explicit features against the other activities.

We further compared the average recognition in terms of the eight activities among all the four methods. We utilize the ROC curve to demonstrate performances for different activities among our method and other three baseline methods. The comparison result is shown in Fig. 6.

As shown in Fig. 6, all the ROC curves are located upon the diagonal line in the upper left corner, which demonstrates the general positive effects for all the methods in the test scenario. It is obvious that the proposed method achieves a substantial improvement in terms of the AUC value comparing with the other three methods. Specifically, our method achieves the AUC value of 0.95, and the DNN, SVM, RF methods can only get the value of 0.87, 0.74, and 0.90 respectively.

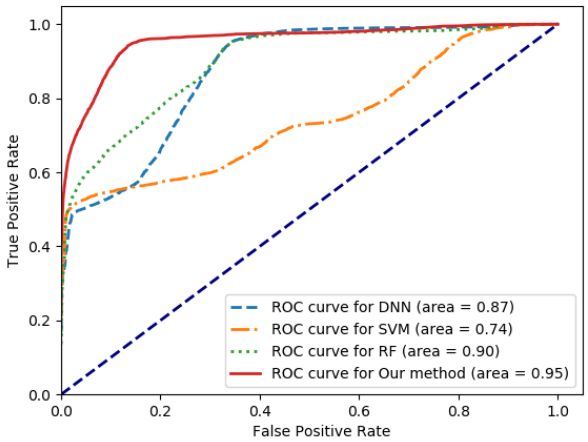


Fig. 6. Comparison result for all Activities

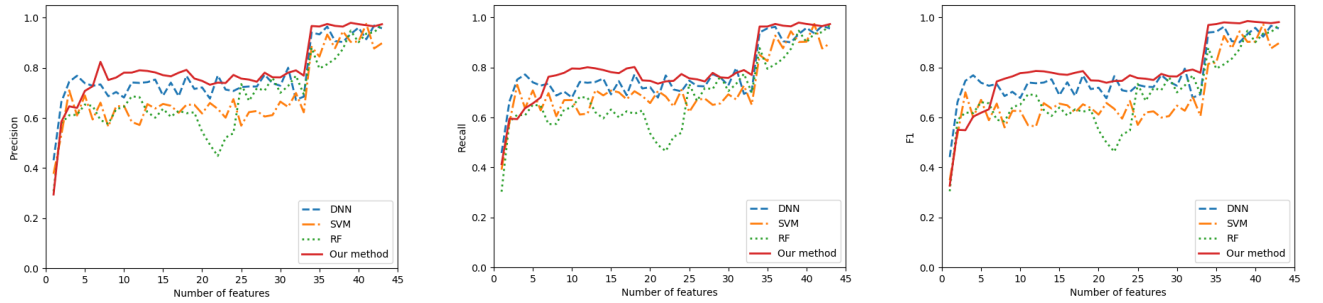
In addition, the body position where the sensor located on would have different influence on the results of activity recognition. To investigate the influence in terms of different sensor positions to the recognition performance, considering results based on the above activity category evaluation, we conducted the evaluation by six different on-body positions (d_1-d_6) for the activity Jumping (A_3). The result illustrated in Table IV depicts the fact that there is no absolute optimal on-body position for activity Jumping when considering all the classifiers. However, according to A_3, the accuracy results of d_1, d_2, d_6 are relatively higher than the others for all the classifiers. It is because these on-body positions can generate more exclusive data to improve the recognition accuracy than from other positions.

TABLE IV
PERFORMANCE OF THE CLASSIFICATION OF DEVICES UNDER DIFFERENT ON-BODY POSITIONS

On-body devices	Our method	DNN	RF	SVM
d_1	**0.96**	**0.95**	**0.94**	**0.89**
d_2	**0.97**	**0.95**	**0.92**	**0.87**
d_3	0.96	0.94	0.92	0.85
d_4	0.95	0.91	0.92	0.78
d_5	0.95	0.92	0.92	0.81
d_6	**0.97**	**0.97**	**0.93**	**0.83**

Moreover, to evaluate how recognition performances of the methods are influenced by the proposed data fusion mechanism, we conducted the experiment by combining different types of sensors data together in terms of Eq. (10). Totally 42 features have been extracted from the subject's profile, accelerometer sensor data, and context sensor data for the evaluation. To investigate how these features affect the methods' recognition performance, we conducted evaluations for all the methods by setting the fused features number from 3 to 42. The experimental results are illustrated in Fig. 7 (a)-(c) respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2020.2985082, IEEE Internet of Things Journal



(a) Precision (b) Recall (c) F1-score

Fig. 7. Comparison results by different number of features. (a) Precision. (b) Recall. (c) F1-score.

The following observations and discussions based on these evaluation results can be noted.

i) The Precision results for the methods are shown in Fig. 7 (a). As shown in the figure, all four precision curves represent general upward trend along with the increasing number of features. Worth to be noted that all the curves arise largely when the number of the features reaches to 33 and become stable at 35, which indicates that the three features (no.33 to no.35) can enhance the performance significantly about 18% and provide important support for HAR in IoHT environments. In addition, although all the methods achieved their best precision results when all the features were included in the evaluation, the proposed method outperforms the others in terms of its maximum precision value and the general performance.

ii) Fig. 7 (b) demonstrates the Recall results for the four methods. Obviously, the neural network based methods (both the DNN and the proposed method) perform better than the conventional machine learning methods (RF and SVM), which suggests that the neural network based methods are relatively suitable to handle this problem. Moreover, the result that our method outperforms the DNN method indicates that the auto-labeling scheme based on semi-supervised learning can benefit the Recall performance. It can partially resolve the inadequately labeled sample situation especially when considering the Recall metric by taking the advantage of leveraging the unlabeled data to enlarge the training set.

iii) At last, we analyze the F1-score metric for the methods as shown in Fig. 7 (c). We observe that the proposed method achieves a substantial improvement in terms of F1-score than the other three methods. The overall F1-score of the proposed method is higher than the others especially when all the features are considered. It indicates that the data fusion upon multi-sensor data can enhance the performance of HAR in IoHT environments.

VI. Conclusion

In this study, we proposed a deep learning enhanced HAR method using weakly labeled sensor data in IoHT environments. Specifically, a semi-supervised deep learning framework was designed to improve the accuracy of HAR in IoHT environments, which consisted of two key function modules as auto-labeling module and LSTM-based classification module. An intelligent auto-labeling scheme was developed based on the DQN technique, which could better solve the problem of inadequately labeled motion data in daily living based on a newly designed distance-based reward rule. A multi-sensor based data fusion mechanism was then developed to integrate the on-body sensor data, context sensor data, and personal profile data together in a seamless way. A LSTM-based neural network was constructed to deal with a series of sequential motion data, and a classification mechanism was improved to identify the fine-grained motion pattern based on the extracted high-level features. Two sensor data sets collected using smartphones and on-body wearable devices were utilized to conduct the experiment. Evaluation results demonstrated the practicability and usefulness of our proposed model and method, comparing with other three baseline methods.

In the future studies, we will go further to study more efficient deep learning techniques for meaningful pattern detection from the weakly labeled data. More evaluations in different situations will be conducted to improve the algorithm with better accuracy and efficiency.

References

[1] Henry Friday Nweke, Ying Wah Teh, Ghulam Mujtaba, Mohammed Ali Al-garadi, “Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions,” Information Fusion, vol. 46, pp. 147-170, 2019.

[2] J. Mills, J. Hu, G. Min, “Communication-Efficient Federated Learning for Wireless Edge Intelligence in IoT,” IEEE Internet of Things Journal, DOI: 10.1109/JIOT.2019.2956615, 2019.

[3] X. Wang, L. T. Yang, H. Liu and M. J. Deen, "A Big Data-as-a-Service Framework: State-of-the-Art and Perspectives," in IEEE Transactions on Big Data, vol. 4, no. 3, pp. 325-340, 1 Sept. 2018.

[4] W. Yang, X. Liu, L. Zhang and L. T. Yang, "Big Data Real-Time Processing Based on Storm," 2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, Melbourne, VIC, 2013, pp. 1784-1787.

[5] O. D. Lara and M. A. Labrador, “A survey on human activity recognition using wearable sensors,” IEEE Commun. Surveys Tuts., vol. 15, no. 3, pp. 1192–1209, 3rd Quart., 2013.

[6] Q. Zhang, L. T. Yang, Z. Yan, Z. Chen and P. Li, "An Efficient Deep Learning Model to Predict Cloud Workload for Industry Informatics," in

IEEE Transactions on Industrial Informatics, vol. 14, no. 7, pp. 3170-3178, July 2018.

[7] P. Turaga, R. Chellappa, et al., "Machine recognition of human activities: A survey.", IEEE Transactions on Circuits and Systems for Video Technology, vol.18, no. 11, pp. 1473-1488, 2008.

[8] M. Daniela, M. Marco, N. Paolo. UniMiB SHAR, "A Dataset for Human Activity Recognition Using Acceleration Data from Smartphones," Applied Sciences, vol. 7, no. 10, pp. 1101-1113, 2017.

[9] Q. Zhu, Z. Chen, C.S. Yeng, "A Novel Semi-Supervised Deep Learning Method for Human Activity Recognition.", IEEE Transactions on Industrial Informatics. vol. 15, no. 7, pp.3821-3830, 2019.

[10] J. He, Q. Zhang, L. Wang and L. Pei, "Weakly Supervised Human Activity Recognition From Wearable Sensors by Recurrent Attention Learning," in IEEE Sensors Journal, vol. 19, no. 6, pp. 2287-2297, 15 March15, 2019.

[11] P. Bharti, D. De, S. Chellappan and S. K. Das, "HuMAn: Complex Activity Recognition with Multi-Modal Multi-Positional Body Sensing," in IEEE Transactions on Mobile Computing, vol. 18, no. 4, pp. 857-870, 1 April 2019.

[12] J. Qi, P. Yang, M. Hanneghan, S. Tang and B. Zhou, "A Hybrid Hierarchical Framework for Gym Physical Activity Recognition and Measurement Using Wearable Sensors," in IEEE Internet of Things Journal, vol. 6, no. 2, pp. 1384-1393, April 2019.

[13] F. Wang, W. Gong and J. Liu, "On Spatial Diversity in WiFi-Based Human Activity Recognition: A Deep Learning-Based Approach," IEEE Internet of Things Journal, vol. 6, no. 2, pp. 2035-2047, 2019.

[14] W. Lu, F. Fan, J. Chu, P. Jing and S. Yuting, "Wearable Computing for Internet of Things: A Discriminant Approach for Human Activity Recognition," IEEE Internet of Things Journal, vol. 6, no. 2, pp. 2749-2759, 2019.

[15] Y. Wang and G. Mori, "Human Action Recognition by Semilatent Topic Models," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 10, pp. 1762-1774, 2009.

[16] Z. Chen, J. Hu, G.Min, A. Zomaya, T. El-Ghazawi, Towards Accurate Prediction for High-Dimensional and Highly-Variable Cloud Workloads with Deep Learning, IEEE Transactions on Parallel and Distributed Systems, DOI: 10.1109/TPDS.2019.2953745, 2019.

[17] N. Krishnan, D. Colbry, C. Juillard, and S. Panchanathan, "Real time human activity recognition using tri-axial accelerometers," in Proc. Sensors, Signals Inf. Process. Workshop, 2008, pp. 3337–3340.

[18] P. Gupta and T. Dallas, "Feature selection and activity recognition system using a single tri-axial accelerometer," IEEE Trans. Biomed. Eng., vol. 61, no. 6, pp. 1780–1786, Jun. 2014.

[19] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in Proc. 17th Conf. Innovative Appl. Artif. Intell., 2005, pp. 1541–1546.

[20] S. W. Lee and K. Mase, "Activity and location recognition using wearable sensors," IEEE Pervasive Comput., vol. 1, no. 3, pp. 24–32, 2002.

[21] J. Mantyjarvi, J. Himberg, and T. Seppanen, "Recognizing human motion with multiple acceleration sensors," in Proc. IEEE Int. Conf. Syst., Man, Cybern., 2001, vol. 2, pp. 747–752.

[22] C. Chen, R. Jafari, N. Kehtarnavaz, A survey of depth and inertial sensor fusion for human action recognition, Multimed. Tools Appl. vol.76, pp. 4405–4425, 2017

[23] L. Bao and S. Intille, "Activity recognition from user-annotated acceleration data," in Proc. Pervasive, 2004, vol. 3001, pp. 1–17.

[24] N. Kern, B. Schiele, H. Junker, P. Lukowicz, and G. Troster, "Wearable sensing to annotate meeting recordings," Pers. Ubiquitous Comput., vol. 7, no. 5, pp. 263–274, 2003.

[25] P. Lukowicz, J. A. Ward, H. Junker, and T. Starner, "Recognizing workshop activity using body worn microphones and accelerometers," in Proc. Pervasive Comput., Apr. 2004, pp. 18–23.

[26] D. J. Patterson, L. Liao, D. Fox, and H. Kautz, "Inferring high-level behavior from low-level sensors," in Proc. 5th Conf. Ubiquitous Comput., 2003, pp. 73–89.

[27] L. Liao, D. J. Patterson, D. Fox, and H. Kautz, "Learning and inferring transporation routines," Artif. Intell., vol. 171, no. 5–6, pp. 311–331, 2007.

[28] M. Liet al., "Multimodal physical activity recognition by fusing temporal and cepstral information," IEEE Trans. Neural Syst. Reha-bil. Eng., vol. 18, no. 4, pp. 369–380, 2010.

[29] Y. Xu, J. Ren, G. Wang, C. Zhang, J. Yang, and Y. Zhang, "A Blockchain-based Nonrepudiation Network Computing Service Scheme for Industrial IoT," IEEE Transactions on Industrial Informatics, vol. 15, no. 6, pp: 3632-3641, 2019.

[30] Y. Xu, J. Ren, Y. Zhang, C. Zhang, B. Shen and Y. Zhang, "Blockchain Empowered Arbitrable Data Auditing Scheme for Network Storage as a Service," IEEE Transactions on Services Computing, DOI: https://doi.org/10.1109/TSC.2019.2953033

[31] O. Chapelle, B. Schcolkopf, and A. Zien, Semi-Supervised Learning. Cambridge, MA, USA: MIT Press, 2006.

[32] A. Blum and T. Mitchell, "Combining labeled and unlabeled data with co-training," in Proc. Annu. Conf. Comput. Learn. Theory, 1998, pp. 92–100.

[33] Z. H. Zhou and M. Li, "Semi-supervised learning by dis-agreement," Knowl. Inf. Syst., vol. 24, no. 3, pp. 415–439, 2010.

[34] M. Stikic, D. Larlus, S. Ebert, and B. Schiele, "Weakly supervised recognition of daily life activities with wearable sensors," IEEE Trans. Pattern Anal. Mach. Intel., vol. 33, no. 12, pp. 2521–2537, Dec. 2011.

[35] Y. Miao, M. Gowayyed, F. Metze, "EESEN: End-to-end speech recognition using deep RNN models and WFST-based decoding", in Proc. Automatic Speech Recognition & Understanding. 2016.

[36] T. Sztyler, H. Stuckenschmidt, W. Petrich. "Position-aware activity recognition with wearable devices," Pervasive and Mobile Computing, vol.38, pp. 281–295, 2017.

Xiaokang Zhou (M'12) is currently an associate professor with the Faculty of Data Science, Shiga University, Japan. He received the Ph.D. degree in human sciences from Waseda University, Japan, in 2014. From 2012 to 2015, he was a research associate with the Department of Human Informatics and Cognitive Sciences, Faculty of Human Sciences, Waseda University, Japan. He also works as a visiting researcher in the RIKEN Center for Advanced Intelligence Project (AIP), RIKEN, Japan, from 2017. Dr. Zhou has been engaged in interdisciplinary research works in the fields of computer science and engineering, information systems, and social and human informatics. His recent research interests include ubiquitous computing, big data, machine learning, behavior and cognitive informatics, cyber-physical-social-system, cyber intelligence and cyber-enabled applications. Dr. Zhou is a member of the IEEE CS, and ACM, USA, IPSJ, Japan, and JSAI, Japan, and CCF, China.

Wei Liang (M'19) received his M.S. and Ph.D. degrees in Computer Science from Central South University in 2005 and 2016. From 2005 to 2012, he worked in Microsoft (China) for soft engineering. From 2014 to 2015, he worked as an exchange researcher in the Department of Human Informatics and Cognitive Sciences, Faculty of Human Sciences, Waseda University, Japan. He is currently working at Key Laboratory of Hunan Province for New Retail Virtual Reality Technology, Hunan University of Technology and Business, China. His research interests include information retrieval, data mining, and artificial intelligence. He has published more than 20 papers at various conferences and journals, including FGCS, JCSS, and PUC. Dr. Liang is a member of the IEEE CS.

Kevin I-Kai Wang (M'04) received the Bachelor of Engineering (Hons.) degree in Computer Systems Engineering and PhD degree in Electrical and Electronics Engineering from the Department of Electrical and Computer Engineering, the University of Auckland, New Zealand, in 2004 and 2009 respectively. He is currently a Senior Lecturer in the Department of Electrical and Computer Engineering, the University of Auckland. He was also a research engineer designing commercial home automation systems and traffic sensing systems from 2009 to 2011. His current research interests include wireless sensor network based ambient intelligence, pervasive healthcare systems, human activity recognition, behavior data analytics and bio-cybernetic systems.

Hao Wang (M'07) is an associate professor in the Department of Computer Science in Norwegian University of Science & Technology, Norway. He has a Ph.D. degree and a B.Eng. degree, both in computer science and engineering, from South China University of Technology. His research interests include big data analytics, industrial internet of things, high performance computing, safety-critical systems, and communication security. He has published 100+ papers in reputable international journals and conferences. He served as a TPC co-chair for IEEE DataCom 2015, IEEE CIT 2017, ES 2017, a senior TPC member for CIKM 2019, and reviewers for journals such as IEEE TKDE, TII, TBD, TETC, T-IFS, IoTJ, TCSS, and ACM TOMM, TIST. He is a member of IEEE IES Technical Committee on Industrial Informatics.

Laurence T. Yang (M'97, SM'15, F'20) received the B.E degree in Computer Science and Technology and B. Sc degree in Applied Physics both from Tsinghua University, China, and the Ph.D. degree in Computer Science from the University of Victoria, Canada. He is a professor and W.F. James Research Chair in the Department of Computer Science, St. Francis Xavier University, Canada. His research interests include parallel and distributed computing, embedded and ubiquitous/pervasive computing, and big data. His research has been supported by the National Sciences and Engineering Research Council, Canada, and the Canada Foundation for Innovation.

Qun Jin (M'95, SM'17) is a professor at the Networked Information Systems Laboratory, Department of Human Informatics and Cognitive Sciences, Faculty of Human Sciences, Waseda University, Japan. He has been extensively engaged in research works in the fields of computer science, information systems, and social and human informatics. He seeks to exploit the rich interdependence between theory and practice in his work with interdisciplinary and integrated approaches. His recent research interests cover human-centric ubiquitous computing, behavior and cognitive informatics, big data, data quality assurance and sustainable use, personal analytics and individual modeling, intelligence computing, blockchain, cyber security, cyber-enabled applications in healthcare, and computing for well-being. Dr. Jin is a senior member of Association of Computing Machinery (ACM), Institute of Electrical and Electronics Engineers (IEEE), and Information Processing Society of Japan (IPSJ).