



ELSEVIER

Contents lists available at ScienceDirect

## Pattern Recognition Letters

journal homepage: [www.elsevier.com/locate/patrec](http://www.elsevier.com/locate/patrec)

## Deep learning for sensor-based activity recognition: A Survey

Jindong Wang<sup>a,b</sup>, Yiqiang Chen<sup>a,b,\*</sup>, Shuji Hao<sup>c</sup>, Xiaohui Peng<sup>a,b</sup>, Lisha Hu<sup>a,b</sup><sup>a</sup> Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China<sup>b</sup> University of Chinese Academy of Sciences, Beijing, China<sup>c</sup> Institute of High Performance Computing, A\*STAR, Singapore

## ARTICLE INFO

## Article history:

Available online xxx

## Keywords:

Deep learning  
Activity recognition  
Pattern recognition  
Pervasive computing

## ABSTRACT

Sensor-based activity recognition seeks the profound high-level knowledge about human activities from multitudes of low-level sensor readings. Conventional pattern recognition approaches have made tremendous progress in the past years. However, those methods often heavily rely on heuristic hand-crafted feature extraction, which could hinder their generalization performance. Additionally, existing methods are undermined for unsupervised and incremental learning tasks. Recently, the recent advancement of deep learning makes it possible to perform automatic high-level feature extraction thus achieves promising performance in many areas. Since then, deep learning based methods have been widely adopted for the sensor-based activity recognition tasks. This paper surveys the recent advance of deep learning based sensor-based activity recognition. We summarize existing literature from three aspects: sensor modality, deep model, and application. We also present detailed insights on existing work and propose grand challenges for future research.

© 2018 Elsevier B.V. All rights reserved.

## 1. Introduction

Human activity recognition (HAR) plays an important role in people's daily life for its competence in learning profound high-level knowledge about human activity from raw sensor inputs. Successful HAR applications include home behavior analysis [61], video surveillance [50], gait analysis [21], and gesture recognition [31]. There are mainly two types of HAR: *video-based* HAR and *sensor-based* HAR [13]. Video-based HAR analyzes videos or images containing human motions from the camera, while sensor-based HAR focuses on the motion data from smart sensors such as an accelerometer, gyroscope, Bluetooth, sound sensors and so on. Due to the thriving development of sensor technology and pervasive computing, sensor-based HAR is becoming more popular and widely used with privacy well protected. Therefore, in this paper, our main focus is on sensor-based HAR.

HAR can be treated as a typical pattern recognition (PR) problem. Conventional PR approaches have made tremendous progress on HAR by adopting machine learning algorithms such as decision tree, support vector machine, naive Bayes, and hidden Markov models [34]. It is no wonder that in some controlled environments where there are only a few labeled data or certain domain knowledge is required (e.g. some disease issues), conventional PR meth-

ods are fully capable of achieving satisfying results. However, in most daily HAR tasks, those methods may heavily rely on heuristic hand-crafted feature extraction, which is usually limited by human domain knowledge [5]. Furthermore, only shallow features can be learned by those approaches [66], leading to undermined performance for unsupervised and incremental tasks. Due to those limitations, the performances of conventional PR methods are restricted regarding classification accuracy and model generalization.

Recent years have witnessed the fast development and advancement of deep learning, which achieves unparalleled performance in many areas such as visual object recognition, natural language processing, and logic reasoning [35]. Different from traditional PR methods, deep learning can largely relieve the effort on designing features and can learn much more high-level and meaningful features by training an end-to-end neural network. In addition, the deep network structure is more feasible to perform unsupervised and incremental learning. Therefore, deep learning is an ideal approach for HAR and has been widely explored in existing work [3,33,47].

Although some surveys have been conducted in deep learning [5,35,58] and HAR [7,34], respectively, there has been no specific survey focusing on the intersections of these two areas. To our best knowledge, this is the *first* article to present the recent advance on deep learning based HAR. We hope this survey can provide a helpful summary of existing work, and present potential future research directions.

\* Corresponding author.

E-mail address: [yqchen@ict.ac.cn](mailto:yqchen@ict.ac.cn) (Y. Chen).<https://doi.org/10.1016/j.patrec.2018.02.010>

0167-8655/© 2018 Elsevier B.V. All rights reserved.

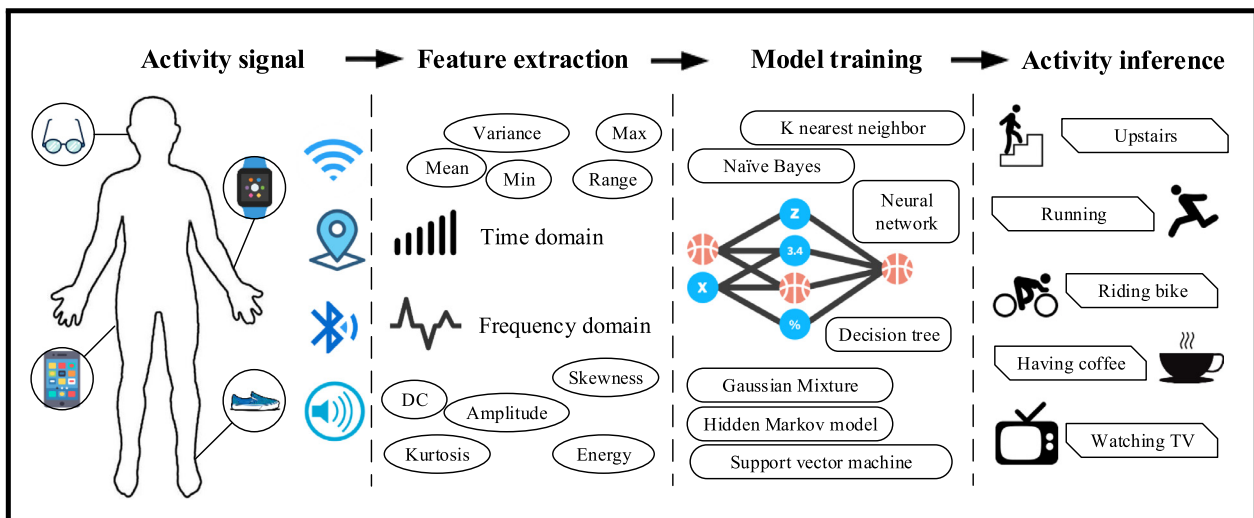


Fig. 1. An illustration of sensor-based activity recognition using conventional pattern recognition approaches.

The rest of this paper is organized as follows. In Section 2, we briefly introduce sensor-based activity recognition and explain why deep learning can improve its performance. In Sections 3, 4 and 5, we review recent advance of deep learning based HAR from three aspects: sensor modality, deep model, and application, respectively. We also introduce several benchmark datasets. Section 6 presents summary and insights on existing work. In Section 7, we discuss some grand challenges and feasible solutions. Finally, this paper is concluded in Section 8.

## 2. Background

### 2.1. Sensor-based activity recognition

HAR aims to understand human behaviors which enable the computing systems to proactively assist users based on their requirement [7]. Formally speaking, suppose a user is performing some kinds of activities belonging to a predefined activity set  $A$ :

$$A = \{A_i\}_{i=1}^m \quad (1)$$

where  $m$  denotes the number of activity types. There is a sequence of sensor reading that captures the activity information

$$\mathbf{s} = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_t, \dots, \mathbf{d}_n\} \quad (2)$$

where  $\mathbf{d}_t$  denotes the sensor reading at time  $t$ .

We need to build a model  $\mathcal{F}$  to predict the activity sequence based on sensor reading  $\mathbf{s}$

$$\hat{A} = \{\hat{A}_j\}_{j=1}^n = \mathcal{F}(\mathbf{s}), \quad \hat{A}_j \in A \quad (3)$$

while the true activity sequence (ground truth) is denoted as

$$A^* = \{A_j^*\}_{j=1}^n, \quad A_j^* \in A \quad (4)$$

where  $n$  denotes the length of sequence and  $n \geq m$ .

The goal of HAR is to learn the model  $\mathcal{F}$  by minimizing the discrepancy between predicted activity  $\hat{A}$  and the ground truth activity  $A^*$ . Typically, a positive loss function  $\mathcal{L}(\mathcal{F}(\mathbf{s}), A^*)$  is constructed to reflect their discrepancy.  $\mathcal{F}$  usually does not directly take  $\mathbf{s}$  as input, and it usually assumes that there is a projection function  $\Phi$  that projects the sensor reading data  $\mathbf{d}_i \in \mathbf{s}$  to a  $d$ -dimensional feature vector  $\Phi(\mathbf{d}_i) \in \mathbb{R}^d$ . To that end, the goal turns into minimizing the loss function  $\mathcal{L}(\mathcal{F}(\Phi(\mathbf{d}_i)), A^*)$ .

Fig. 1 presents a typical flowchart of HAR using conventional PR approaches. First, raw signal inputs are obtained from several

types of sensors (smartphones, watches, Wi-Fi, Bluetooth, sound etc.). Second, features are manually extracted from those readings based on human knowledge [4], such as the *mean*, *variance*, *DC*, and *amplitude* in traditional machine learning approaches [25]. Finally, those features serve as inputs to train a PR model to make activity inference in real HAR tasks.

### 2.2. Why deep learning?

Conventional PR approaches have made tremendous progress in HAR [7]. However, there are several drawbacks to conventional PR methods.

Firstly, the features are always extracted via a heuristic and hand-crafted way, which heavily relies on human experience or domain knowledge. This human knowledge may help in certain task-specific settings, but for more general environments and tasks, this will result in a lower chance and longer time to build a successful activity recognition system.

Secondly, only shallow features can be learned according to human expertise [66]. Those shallow features often refer to some statistical information including mean, variance, frequency and amplitude etc. They can only be used to recognize low-level activities like *walking* or *running*, and hard to infer high-level or context-aware activities [67]. For instance, *having coffee* is more complex and nearly impossible to be recognized by using only shallow features.

Thirdly, conventional PR approaches often require a large amount of well-labeled data to train the model. However, most of the activity data are remaining unlabeled in real applications. Thus, these models' performance is undermined in unsupervised learning tasks [5]. In contrast, existing deep generative networks [24] are able to exploit the unlabeled samples for model training.

Moreover, most existing PR models mainly focus on learning from static data; while activity data in real life are coming in stream, requiring robust online and incremental learning.

Deep learning tends to overcome those limitations. Fig. 2 shows how deep learning works for HAR with different types of networks. Compared to Fig. 1, the feature extraction and model building procedures are often performed simultaneously in the deep learning models. The features can be learned automatically through the network instead of being manually designed. Besides, the deep neural network can also extract high-level representation in deep layer, which makes it more suitable for complex activity recognition

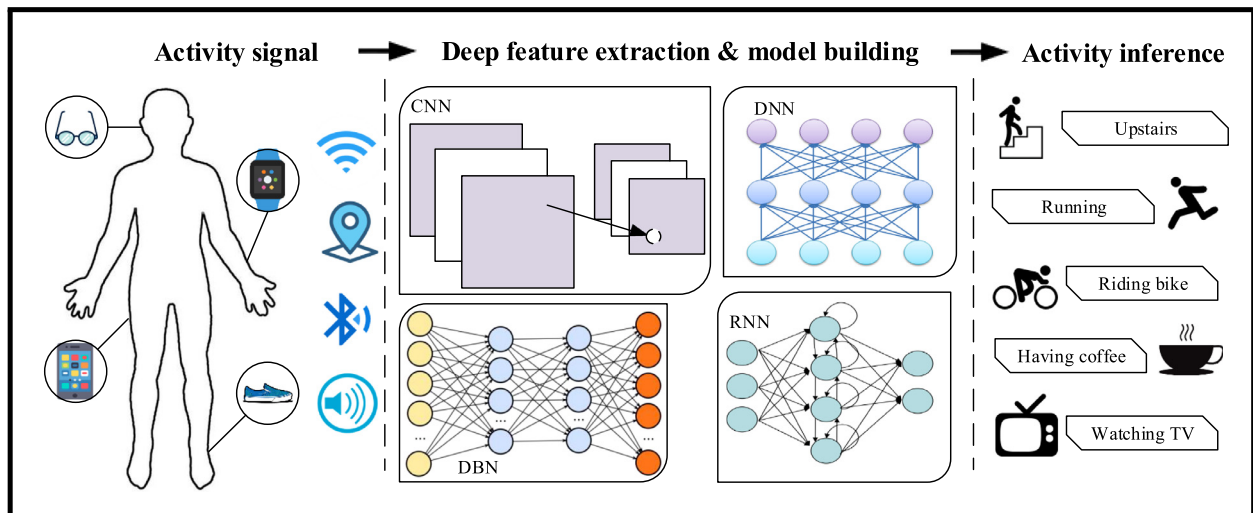


Fig. 2. An illustration of sensor-based activity recognition using deep learning approaches.

**Table 1**  
Sensor modalities for HAR tasks.

Modality	Description	Examples
Body-worn	Worn by the user to describe the body movements	Smartphone, watch, or band's accelerometer, gyroscope etc.
Object	Attached to objects to capture objects movements	RFID, accelerometer on cup etc.
Ambient	Applied in environment to reflect user interaction	Sound, door sensor, WiFi, Bluetooth etc.
Hybrid	Crossing sensor boundary	Combination of types, often deployed in smart environments

tasks. When faced with a large amount of unlabeled data, deep generative models [24] are able to exploit the unlabeled data for model training. What's more, deep learning models trained on a large-scale labeled dataset can usually be transferred to new tasks where there are few or none labels. In the following sections, we mainly summarize the existing work based on the pipeline of HAR: (a) *sensor modality*, (b) *deep model*, and (c) *application*.

### 3. Sensor modality

Although some HAR approaches can be generalized to all sensor modalities, most of them are only specific to certain types. According to [8], we mainly classify those modalities into three aspects: *body-worn sensors*, *object sensors*, and *ambient sensors*. Table 1 briefly outlines all the modalities.

#### 3.1. Body-worn sensor

Body-worn sensors are one of the most common modalities in HAR. Those sensors are often worn by the users, such as an accelerometer, magnetometer, and gyroscope. The acceleration and angular velocity are changed according to human body movements; thus they can infer human activities. Those sensors can often be found on smart phones, watches, bands, glasses, and helmets.

Body-worn sensors were widely used in deep learning based HAR [10,27,47,66,70]. Among those work, the accelerometer is mostly adopted. Gyroscope and magnetometer are also frequently used together with the accelerometer. Those sensors are often exploited to recognize activities of daily living (ADL) and sports. Instead of extracting statistical and frequency features from the movement data, the original signal is directly used as inputs for the network.

#### 3.2. Object sensor

Object sensors are usually placed on objects to detect the movement of a specific object [8]. Different from body-worn sensors which capture human movements, object sensors are mainly used to detect the movement of certain objects in order to infer human activities. For instance, the accelerometer attached to a cup can be used to detect the *drinking water* activity. Radio frequency identifier (RFID) tags are typically used as object sensors and deployed in smart home environment [15,61,66] and medical activities [38,63]. The RFID can provide more fine-grained information for more complex activity recognition.

It should be noted that object sensors are less used than body-worn sensors due to the difficulty in its deployment. Besides, the combination of object sensors with other types is emerging in order to recognize more high-level activities [67].

#### 3.3. Ambient sensor

Ambient sensors are used to capture the interaction between humans and the environment. They are usually embedded in users' smart environment. There are many kinds of ambient sensors such as radar, sound sensors, pressure sensors, and temperature sensors. Different from object sensors which measure the object movements, ambient sensors are used to capture the change of the environment.

Several literature used ambient sensors to recognize daily activities and hand gesture [31,33,63]. Most of the work was tested in the smart home environment. Same as object sensors, the deployment of ambient sensors is also difficult. In addition, ambient sensors are easily affected by the environment, and only certain types of activities can be robustly inferred.

**Table 2**  
Deep learning models for HAR tasks.

Model	Description
DNN	Deep fully-connected network, artificial neural network with deep layers
CNN	Convolutional neural network, multiple convolution operations for feature extraction
RNN	Recurrent neural network, network with time correlations and LSTM
DBN/RBM	Deep belief network and restricted Boltzmann machine
SAE	Stacked autoencoder, feature learning by decoding-encoding autoencoder
Hybrid	combination of some deep models

### 3.4. Hybrid sensor

Some work combined different types of sensors for HAR. As shown in [23], combining acceleration with acoustic information could improve the accuracy of HAR. Ambient sensors are also used together with object sensors; hence they can record both the object movements and environment state. [61] designed a smart home environment called *A-Wristocracy*, where a large number of fine-grained and complex activities of multiple occupants can be recognized through body-worn, object, and ambient sensors. It is obvious that the combination of sensors is capable of capturing rich information of human activities, which is also possible for a real smart home system in the future.

## 4. Deep model

In this section, we investigate the deep learning models used in HAR tasks. Table 2 lists all the models.

### 4.1. Deep neural network

Deep neural network (DNN) is developed from artificial neural network (ANN). Traditional ANN often contains very few hidden layers (shallow) while DNN contains more (deep). With more layers, DNN is more capable of learning from large data. DNN usually serves as the dense layer of other deep models. For example, in a convolution neural network, several dense layers are often added after the convolution layers. In this part, we mainly focus on DNN as a single model, while in other sections we will discuss the dense layer.

[61] first extracted hand-engineered features from the sensors, then those features are fed into a DNN model. Similarly, [62] performed PCA before using DNN. In those work, DNN only served as a classification model after hand-crafted feature extraction, hence they may not generalize well. And the network was rather shallow. [21] used a 5-hidden-layer DNN to perform automatic feature learning and classification with improved performance. Those work indicated that, when the HAR data is multi-dimensional and activities are more complex, more hidden layers can help the model train well since their representation capability is stronger [5]. However, more details should be considered in certain situations to help the model fine-tune better.

### 4.2. Convolutional neural network

Convolutional Neural Network (ConvNets, or CNN) leverages three important ideas: sparse interactions, parameter sharing, and equivariant representations [35]. After convolution, there are usually pooling and fully-connected layers, which perform classification or regression tasks.

CNN is competent to extract features from signals and it has achieved promising results in image classification, speech recognition, and text analysis. When applied to time series classification like HAR, CNN has two advantages over other models: local dependency and scale invariance. Local dependency means the nearby

signals in HAR are likely to be correlated, while scale invariance refers to the scale-invariant for different paces or frequencies. Due to the effectiveness of CNN, most of the surveyed work focused on this area.

When applying CNN to HAR, there are several aspects to be considered: *input adaptation*, *pooling*, and *weight-sharing*.

1) *Input adaptation*. Unlike images, most HAR sensors produce *time series* readings such as acceleration signal, which is temporal multi-dimensional 1D readings. Input adaptation is necessary before applying CNN to those inputs. The main idea is to adapt the inputs in order to form a *virtual image*. There are mainly two types of adaptation: *model-driven* and *data-driven*.

- *Data-driven* approach treats each dimension as a channel, then performs 1D convolution on them. After convolution and pooling, the outputs of each channel are flattened to unified DNN layers. A very early work is [70], where each dimension of the accelerometer was treated as one channel like RGB of an image, then the convolution and pooling were performed separately. [66] further proposed to unify and share weights in multi-sensor CNN by using 1D convolution in the same temporal window. Along with this line, [10] resized the convolution kernel to obtain the best kernel for HAR data. Other similar work include [21,48,57]. This data-driven approach treats the 1D sensor reading as a 1D image, which is simple and easy to implement. The disadvantage of this approach is the ignorance of dependencies between dimension and sensors, which may influence the performance.
- *Model-driven* approach resizes the inputs to a virtual 2D image so as to adopt a 2D convolution. This approach usually pertains to non-trivial input tuning techniques. [19] combined all dimensions to form an image, while [27] designed a more complex algorithm to transform the time series into an image. In [59], pressure sensor data was transformed to the image via modality transformation. Other similar work include [38,52]. This model-driven approach can make use of the temporal correlation of sensor. But the map of time series to image is non-trivial task and needs domain knowledge.

2) *Pooling*. The *convolution-pooling* combination is common in CNN, and most approaches performed max or average pooling after convolution [19,31,48]. Apart from avoiding overfitting, pooling can also speed up the training process on large data [5].

3) *Weight-sharing*. Weight sharing [57,69] is an efficient method to speed up the training process on a new task. [70] utilized a relaxed partial weight sharing technique since the signal appeared in different units may behave differently. [18] adopted a CNN-pf and CNN-pff structure to investigate the performance of different weight-sharing techniques. It is shown in those literature that partial weight-sharing could improve the performance of CNN.

### 4.3. Autoencoder

Autoencoder learns a latent representation of the input values through the hidden layers, which can be considered as an encoding-decoding procedure. The purpose of autoencoder is to

learn more advanced feature representation via an unsupervised learning schema. Stacked autoencoder (SAE) is the stack of some autoencoders. SAE treats every layer as the basic model of autoencoder. After several rounds of training, the learned features are stacked with labels to form a classifier.

[2,63] used SAE for HAR, where they first adopted the greedy layer-wise pre-training [24], then performed fine-tuning. Compared to those works, [39] investigated the sparse autoencoder by adding KL divergence and noise to the cost function, which indicates that adding sparse constraints could improve the performance of HAR. The advantage of SAE is that it can perform unsupervised feature learning for HAR, which could be a powerful tool for feature extraction. But SAE depends too much on its layers and activation functions which may be hard to search the optimal solutions.

#### 4.4. Restricted Boltzmann machine

Restricted Boltzmann machine (RBM) is a bipartite, fully-connected, undirected graph consisting of a visible layer and a hidden layer [24]. The stacked RBM is called deep belief network (DBN) by treating every two consecutive layers as an RBM. DBN/RBM is often followed by fully-connected layers.

In pre-training, most work applied Gaussian RBM in the first layer while binary RBM for the rest layers [20,33,47]. For multi-modal sensors, [51] designed a multi-modal RBM where an RBM is constructed for each sensor modality, then the output of all the modalities are unified. [37] added pooling after the fully-connected layers to extract the important features. [15] used a contrastive gradient (CG) method to update the weight in fine-tuning, which helps the network to search and convergence quickly in all directions. [72] further implemented RBM on a mobile phone for offline training, indicating RBM can be very light-weight. Similar to autoencoder, RBM/DBN can also perform unsupervised feature learning for HAR.

#### 4.5. Recurrent neural network

Recurrent neural network (RNN) is widely used in speech recognition and natural language processing by utilizing the temporal correlations between neurons. LSTM (long-short term memory) cells are often combined with RNN where LSTM is serving as the *memory* units through gradient descent.

Few work used RNN for the HAR tasks [14,17,21,26], where the learning speed and resource consumption are the main concerns for HAR. [26] investigated several model parameters first and then proposed a relatively *good* model which can perform HAR with high throughput. [14] proposed a binarized-BLSTM-RNN model, in which the weight parameters, input, and output of all hidden layers are all binary values. The main line of RNN based HAR models is dealing with resource-constrained environments while still achieve good performance.

#### 4.6. Hybrid model

Hybrid model is the combination of some deep models.

One emerging hybrid model is the combination of CNN and RNN. [44,68] provided good examples for how to combine CNN and RNN. It is shown in [44] that the performance of ‘CNN + recurrent dense layers’ is better than ‘CNN + dense layers’. Similar results are also shown in [59]. The reason is that CNN is able to capture the spatial relationship, while RNN can make use of the temporal relationship. Combining CNN and RNN could enhance the ability to recognize different activities that have varied time span and signal distributions. Other work combined CNN with models such as SAE [77] and RBM [40]. In those work, CNN performs feature extraction, and the generative models can help in speeding up

the training process. In the future, we expect there will be more research in this area.

## 5. Applications

HAR is always not the final goal of an application, but it serves as an important step in many applications such as skill assessment and smart home assistant. In this section, we survey deep learning based HAR from the application perspective.

### 5.1. Featured applications

Most of the surveyed work focused on recognizing *activities of daily living* (ADL) and *sports* [10,53,56,70]. Those activities of simple movements are easily captured by body-worn sensors. Some research studied people’s *lifestyle* such as sleep [57] and respiration [22,29]. The detection of such activities often requires some object and ambient sensors such as WiFi and sound, which are rather different from ADL.

It is a developing trend to apply HAR to *health and disease* issues. Some pioneering work has been done with regard to Parkinson’s disease [20], trauma resuscitation [37,38] and paroxysmal atrial fibrillation (PAF) [48]. Disease issues are always related to the change of certain body movements or functions, so they can be detected using corresponding sensors.

Under those circumstances, the association between disease and activity should be given more consideration. It is important to use the appropriate sensors. For instance, Parkinson’s disease is often related to the frozen of gait, which can be reflected by some inertial sensors attached to shoes [20].

Other than health and disease, the recognition of *high-level* activities is helpful to learn more resourceful information for HAR. The movement, behavior, environment, emotion, and thought are critical parts in recognizing high-level activities. However, most work only focused on body movements in smart homes [15,61], which is not enough to recognize high-level activities. For instance, [61] combined activity and environment signal to recognize activities in a smart home, but the activities are constrained to body movements without more information on user emotion and state, which are also important. In the future, we expect there will be more research in this area.

### 5.2. Benchmark datasets

We extensively explore the benchmark datasets for deep learning based HAR. Basically, there are two types of data acquisition schemes: *self data collection* and *public datasets*.

- *Self data collection*: Some work performed their own data collection (e.g. [6,10,71,72]). Very detailed efforts are required for self data collection, and it is rather tedious to process the collected data.
- *Public datasets*: There are already many public HAR datasets that are adopted by most researchers (e.g. [21,47,52]). By summarizing existing literature, we present several widely used public datasets in Table 3.

## 6. Summary and discussion

Table 4 presents all the surveyed work in this article. We can make several observations based on the table.

**1) Sensor deployment and preprocessing.** Choosing the suitable sensors is critical for successful HAR. In surveyed literature, body-worn sensors serve as the most common modalities and accelerometer is mostly used. The reasons are two folds. Firstly, a lot of wearable devices such as smartphones or watches are equipped

**Table 3**

Public HAR datasets (A=accelerometer, G=gyroscope, M=magnetometer, O=object sensor, AM=ambient sensor, ECG=electrocardiograph).

ID	Dataset	Type	#Subject	S. Rate	#Activity	#Sample	Sensor	Reference
D01	OPPORTUNITY	ADL	4	32 Hz	16	701,366	A, G, M, O, AM	[44]
D02	Skoda Checkpoint	Factory	1	96 Hz	10	22,000	A	[47]
D03	UCI Smartphone	ADL	30	50 Hz	6	10,299	A, G	[2]
D04	PAMAP2	ADL	9	100 Hz	18	2,844,868	A, G, M	[76]
D05	USC-HAD	ADL	14	100 Hz	12	2,520,000	A, G	[27]
D06	WISDM	ADL	29	20 Hz	6	1,098,207	A	[3]
D07	DSADS	ADL	8	25 Hz	19	1,140,000	A, G, M	[73]
D08	Ambient kitchen	Food preparation	20	40 Hz	2	55,000	O	[47]
D09	Darmstadt Daily Routines	ADL	1	100 Hz	35	24,000	A	[47]
D10	Actitracker	ADL	36	20 Hz	6	2,980,765	A	[70]
D11	SHO	ADL	10	50 Hz	7	630,000	A, G, M	[27]
D12	BIDMC	Heart failure	15	125 Hz	2	>20,000	ECG	[76]
D13	MHEALTH	ADL	10	50 Hz	12	16,740	A, C, G	[18]
D14	Daphnet Gait	Gait	10	64 Hz	2	1,917,887	A	[21]
D15	ActiveMiles	ADL	10	50–200 Hz	7	4,390,726	A	[53]
D16	HASC	ADL	1	200 Hz	13	NA	A	[23]
D17	PAF	PAF	48	128 Hz	2	230,400	EEG	[48]
D18	ActRecTut	Gesture	2	32 Hz	12	102,613	A, G	[66]
D19	Heterogeneous	ADL	9	100–200 Hz	6	43,930,257	A, G	[68]

with an accelerometer, which is easy to access. Secondly, the accelerometer is competent to recognize many types of daily activities since most of them are simple body movements. Compared to body-worn sensors, object and ambient sensors are better at recognizing activities related to context and environment such as *having coffee*. Therefore, it is suggested to use body-worn sensors (mostly accelerometer+gyroscope) for ADL and sports activities. If the activities are pertaining to some semantic meaning but more than simple body movements, it is better to combine the object and ambient sensors. In addition, there are few public datasets for object and ambient sensors probably because of privacy issues and deployment difficulty of the data collecting system. We expect there will be more open datasets regarding those sensors.

Sensor placement is also important. Most body-worn sensors are placed on the dominant wrist, waist, and the dominant hip pocket. This placement strategy can help to recognize most common daily activities. However, when it comes to object and ambient sensors, it is critical to deploy them in a non-invasive way. Those sensors are not usually interacting with users directly, so it is critical to collect the data naturally and non-invasively.

Before using deep models, the raw sensor data need to be pre-processed accordingly. There are two important aspects. The first aspect is *sliding window*. The inputs should be cut into individual inputs according to the sampling rate. This procedure is similar to conventional PR approaches. The second one is *channels*. Different sensor modalities can be treated as separate channels, and each axis of a sensor can also be a channel. Using multi-channel could enhance the representation capability of the deep model since it can reflect the hidden knowledge of the sensor inputs.

**2) Model selection.** There are several deep models surveyed in this article. Then, a natural question arises: *which model is the best for HAR?* [21] did an early work by investigating the performance of DNN, CNN and RNN through 4000 experiments on some public HAR datasets. We combine their work and our explorations to draw some conclusions: RNN and LSTM are recommended to recognize short activities that have natural order while CNN is better at inferring long-term repetitive activities [21]. The reason is that RNN could make use of the time-order relationship between sensor readings, and CNN is more capable of learning deep features contained in recursive patterns. For multi-modal signals, it is better to use CNN since the features can be integrated through multi-channel convolutions [19,70,76]. While adapting CNN, data-driven approaches are better than model-driven approaches as the inner properties of the activity signal can be exploited better when the

input data are transformed into the virtual image. Multiple convolutions and poolings also help CNN perform better. RBM and autoencoders are usually pre-trained before being fine-tuned. Multi-layer RBM or SAE is preferred for more accurate recognition.

Technically there is no model which outperforms all the others in all situations, so it is recommended to choose models based on the scenarios. To better illustrate the performance of some deep models, Table 5 offers some results comparison of existing work on public datasets in Table 3<sup>1</sup> In Skoda and UCI Smartphone protocols, CNN achieves the best performance. In two OPPORTUNITY protocols, DBN and RNN outperform the others. This confirms that no models can achieve the best in all tasks. Moreover, the hybrid models tend to perform better than single models (DeepConvLSTM in OPPORTUNITY 1 and Skoda). For a single model, CNN with shifted inputs (Fourier transform) generates better results compared to shifted kernels.

## 7. Grand challenges

Despite the progress in previous work, there are still challenges for deep learning based HAR. In this section, we present those challenges and propose some feasible solutions.

**A. Online and mobile deep activity recognition.** Two critical issues are related to deep HAR: online deployment and mobile application. Although some existing work adopted deep HAR on smartphone [33] and watch [6], they are still far from online and mobile deployment. Because the model is often trained offline on some remote server and the mobile device only utilizes a trained model. This approach is neither real-time nor friendly to incremental learning. There are two approaches to tackle this problem: *reducing the communication cost between mobile and server*, and *enhancing computing ability of the mobile devices*.

**B. More accurate unsupervised activity recognition.** The performance of deep learning still relies heavily on labeled samples. Acquiring sufficient activity labels is expensive and time-consuming. Thus, *unsupervised activity recognition* is urgent.

- *Take advantage of the crowd.* The latest research indicates that exploiting the knowledge from the crowd will facilitate the task [49]. Crowd-sourcing takes advantage of the crowd to annotate the unlabeled activities. Other than acquiring labels pas-

<sup>1</sup> OPP 1, OPP 2, Skoda, and UCI smartphone follow the protocols in [21], [47], [70], and [56], respectively. OPP 1 used weighted f1-score; OPP 2, Skoda, and UCI smartphone used accuracy.

**Table 4**

Summation of existing works based on the three aspects: sensor modality, deep model and application (in literature order).

Literature	Sensor Modality	Deep Model	Application	Dataset
[2]	Body-worn	SAE	ADL	D03
[3]	Body-worn	RBM	ADL, factory, Parkinson	D02, D06, D14
[6]	Body-worn, ambient	RBM	Gesture, ADL, transportation	Self, D01
[10]	Body-worn	CNN	ADL	Self
[11]	Body-worn	CNN	ADL	D06
[12]	Body-worn	DNN	Parkinson	Self
[14]	Body-worn	RNN	ADL	D01, D04, Self
[15]	Object, ambient	DBN	ADL	Self
[16]	Body-worn	CNN	ADL	Self, D01
[17]	Body-worn, object, ambient	RNN	ADL, smart home	D01, D02, D04
[19]	Body-worn	CNN	Factory, health	D02, D13
[18]	Body-worn	CNN	ADL, health	D13
[20]	Body-worn	RBM	Parkinson	Self
[21]	Body-worn, object, ambient	DNN, CNN, RNN	ADL, smart home, gait	D01, D04, D14
[22]	Body-worn	CNN	Gait	Self
[23]	Body-worn, ambient	RBM	ADL, smart home	D16
[26]	Body-worn	RNN	ADL	D16
[27]	Body-worn	CNN	ADL	D03, D05, D11
[29]	Ambient	CNN	Respiration	Self
[31]	Ambient	CNN	Hand gesture	Self
[30]	Body-worn	CNN	ADL	Self
[32]	Body-worn, ambient	RBM	ADL, emotion	Self
[33]	Ambient	RBM	ADL	Self
[36]	Body-worn	CNN	ADL	Self
[37]	Object	RBM	Patient resuscitation	Self
[38]	Object	CNN	Patient resuscitation	Self
[39]	Body-worn	SAE	ADL	D03
[40]	Body-worn	CNN, RBM	ADL	Self
[41]	Body-worn	CNN	ADL, gesture	Self
[42]	Body-worn	CNN	ADL, smart home	D01, D02
[43]	Body-worn	RNN	ADL, smart home	D01, D02, D05, D14
[44]	Body-worn	CNN, RNN	ADL, gesture, posture, factory	D01, D02
[46]	Body-worn	CNN	ADL	Self
[47]	Body-worn, object	RBM	ADL, food preparation, factory	D01, D02, D08, D14
[48]	Body-worn	CNN	PAF disease	D17
[51]	Body-worn	RBM	ADL	D19
[52]	Body-worn	CNN	ADL, factory	D02, D06, D14, D15
[53]	Body-worn	CNN	ADL, factory, Parkinson	D02, D06, D14, D15
[54–56]	Body-worn	CNN	ADL	D03
[57]	Body-worn	CNN, RNN, DNN	ADL, sleep	Self
[59]	Ambient	CNN, RNN	Gait	NA
[61]	Body-worn, object, ambient	DNN	ADL	Self
[62]	Body-worn	DNN	ADL	D03
[65]	Body-worn, ambient	CNN	ADL, location	Self
[63]	Object, ambient	SAE	ADL	NA
[66]	Body-worn, object, ambient	CNN	ADL, smart home, gesture	D01, D18
[68]	Body-worn, object	CNN, RNN	Cartrack, ADL	Self, D19
[69]	Body-worn	CNN	ADL	Self
[70]	Body-worn, ambient, object	CNN	ADL, smart home, factory	D01, D02, D10
[71]	Body-worn	DNN	ADL	Self
[72]	Body-worn	RBM	ADL	Self
[73]	Body-worn	DBN	ADL, smart home	D01, D05, D07
[75]	Object	CNN	Medical	Self
[74]	Body-worn	DNN	ADL	Self
[77]	Body-worn	CNN, SAE	ADL	D04
[76]	Body-worn	CNN	ADL, heart failure	D04, D14

sively, researchers could also develop more elaborate, privacy-concerned way to collect useful labels.

- *Deep transfer learning.* Transfer learning performs data annotation by leveraging labeled data from other auxiliary domains [13,45,64]. There are many factors related to human activity, which can be exploited as auxiliary information using deep transfer learning. Problems such as sharing weights between networks, exploiting knowledge between activity related domains, and how to find more relevant domains are to be resolved.

**C. Flexible models to recognize high-level activities.** More complex high-level activities need to be recognized other than only simple daily activities. It is difficult to determine the hierarchical structure of high-level activities because they contain more seman-

tic and context information. Existing methods often ignore the correlation between signals, thus they cannot obtain good results.

- *Hybrid sensor.* Elaborate information provided by the hybrid sensor is useful for recognizing fine-grained activities [61]. Special attention should be paid to the recognition of fine-grained activities by exploiting the collaboration of hybrid sensors.
- *Exploit context information.* Context is any information that can be used to characterize the situation of an entity [1]. Context information such as Wi-Fi, Bluetooth, and GPS can be used to infer more environmental knowledge about the activity. The exploitation of resourceful context information will greatly help to recognize user state as well as more specific activities.

**Table 5**  
Performance comparison of existing deep models.

Protocol	Model	Result	Reference
OPP 1	b-LSTM-S	<b>92.70</b>	[21]
	CNN	85.10	[66]
	CNN	88.30	[44]
OPP 2	DeepConvLSTM	91.70	[44]
	DBN	73.20	[47]
	CNN	76.80	[70]
Skoda	DBN	<b>83.30</b>	[73]
	CNN	86.10	[70]
	CNN	89.30	[3]
UCI smartphone	DeepConvLSTM	<b>95.80</b>	[44]
	CNN	94.61	[56]
	CNN	<b>95.18</b>	[27]
	CNN	94.79	[54]
	CNN	90.00	[55]

**D. Light-weight deep models.** Deep models often require lots of computing resources, which is not available for wearable devices. In addition, the models are often trained off-line which cannot be executed in real-time. However, less complex models such as shallow NN and conventional PR methods could not achieve good performance. Therefore, it is necessary to develop light-weight deep models to perform HAR.

- *Combination of human-crafted and deep features.* Recent work indicated that human-crafted and deep features together could achieve better performance [47]. Some pre-knowledge about the activity will greatly contribute to more robust feature learning in deep models [60]. Researchers should consider the possibility of applying two kinds of features to HAR with human experience and machine intelligence.
- *Collaboration of deep and shallow models.* Deep models have powerful learning abilities, while shallow models are more efficient. The collaboration of those two models has the potential to perform both accurate and light-weight HAR. Several issues such as how to share the parameters between deep and shallow models are to be addressed.

**E. Non-invasive activity sensing.** Traditional activity collection strategies need to be updated with more non-invasive approaches. Non-invasive approaches tend to collect information and infer activity without disturbing the subjects and requires more flexible computing resources.

- *Opportunistic activity sensing with deep learning.* Opportunistic sensing could dynamically harness the non-continuous activity signal to accomplish activity inference [9]. In this scenario, back propagation of deep models should be well-designed.

**F. Beyond activity recognition: assessment and assistant.** Recognizing activities is often the initial step in many applications. For instance, some professional skill assessment is required in fitness exercises and smart home assistant plays an important role in healthcare services. There is some early work on climbing assessment [28]. With the advancement of deep learning, more applications should be developed to be beyond just recognition.

## 8. Conclusion

Human activity recognition is an important research topic in pattern recognition and pervasive computing. In this paper, we survey the recent advance in deep learning approaches for sensor-based activity recognition. Compared to traditional pattern recognition methods, deep learning reduces the dependency on human-crafted feature extraction and achieves better performance by automatically learning high-level representations of the sensor data. We highlight the recent progress in three important categories:

sensor modality, deep model, and application. Subsequently, we summarize and discuss the surveyed research in detail. Finally, several grand challenges and feasible solutions are presented for future research.

## Acknowledgments

This work is supported in part by National Key Research & Development Program of China (No.2017YFB1002801), NSFC (No.61572471), and Science and Technology Planning Project of Guangdong Province (No.2015B010105001). Authors thank the reviewers for their valuable comments.

## References

- [1] G.D. Abowd, A.K. Dey, P.J. Brown, N. Davies, M. Smith, P. Steggles, Towards a better understanding of context and context-awareness, in: International Symposium on Handheld and Ubiquitous Computing, Springer, 1999, pp. 304–307.
- [2] B. Almasluh, J. AlMuhtadi, A. Artoli, An effective deep autoencoder approach for online smartphone-based human activity recognition, *Int. J. Comput. Sci. Netw. Security* 17 (4) (2017) 160.
- [3] M.A. Alsheikh, A. Selim, D. Niyato, L. Doyle, S. Lin, H.-P. Tan, Deep activity recognition models with triaxial accelerometers, AAAI Workshop (2016).
- [4] L. Bao, S.S. Intille, Activity recognition from user-annotated acceleration data, in: International Conference on Pervasive Computing, Springer, 2004, pp. 1–17.
- [5] Y. Bengio, Deep learning of representations: Looking forward, in: International Conference on Statistical Language and Speech Processing, Springer, 2013, pp. 1–37.
- [6] S. Bhattacharya, N.D. Lane, From smart to deep: Robust activity recognition on smartwatches using deep learning, in: 2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops), IEEE, 2016, pp. 1–6.
- [7] A. Bulling, U. Blanke, B. Schiele, A tutorial on human activity recognition using body-worn inertial sensors, *ACM Comput. Surv.* 46 (3) (2014) 33.
- [8] R. Chavarriaga, H. Sagh, A. Calatroni, S.T. Digumarti, G. Tröster, J.d.R. Millán, D. Roggen, The opportunity challenge: a benchmark database for on-body sensor-based activity recognition, *Pattern Recognit. Lett.* 34 (15) (2013) 2033–2042.
- [9] Y. Chen, Y. Gu, X. Jiang, J. Wang, Ocean: a new opportunistic computing model for wearable activity recognition, in: UbiComp, ACM, 2016, pp. 33–36.
- [10] Y. Chen, Y. Xue, A deep learning approach to human activity recognition based on single accelerometer, in: Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on, IEEE, 2015, pp. 1488–1492.
- [11] Y. Chen, K. Zhong, J. Zhang, Q. Sun, X. Zhao, Lstm networks for mobile human activity recognition (2016b).
- [12] W.-Y. Cheng, A.e.a. Scotland, Human activity recognition from sensor-based large-scale continuous monitoring of parkinsons disease patients, in: Connected Health: Applications, Systems and Engineering Technologies (CHASE), 2017 IEEE/ACM International Conference on, 2017, pp. 249–250.
- [13] D. Cook, K.D. Feuz, N.C. Krishnan, Transfer learning for activity recognition: a survey, *Knowl. Inf. Syst.* 36 (3) (2013) 537–556.
- [14] M. Edel, E. Köppe, Binarized-blstm-rnn based human activity recognition, in: Indoor Positioning and Indoor Navigation (IPIN), 2016 International Conference on, IEEE, 2016, pp. 1–7.
- [15] H. Fang, C. Hu, Recognizing human activity in smart home using deep learning algorithm, in: Chinese Control Conference (CCC), 2014, pp. 4716–4720.
- [16] H. Gjoreski, J. Bizjak, M. Gjoreski, M. Gams, Comparing deep and classical machine learning methods for human activity recognition using wrist accelerometer, IJCAI-16 workshop on Deep Learning for Artificial Intelligence (DLAI), 2016.
- [17] Y. Guan, T. Ploetz, Ensembles of deep lstm learners for activity recognition using wearables, (2017). arXiv preprint arXiv:1703.09370.
- [18] S. Ha, S. Choi, Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors, in: Neural Networks (IJCNN), 2016 International Joint Conference on, IEEE, 2016, pp. 381–388.
- [19] S. Ha, J.-M. Yun, S. Choi, Multi-modal convolutional neural networks for activity recognition, in: Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on, IEEE, 2015, pp. 3017–3022.
- [20] N.Y. Hammerla, J. Fisher, P. Andras, L. Rochester, R. Walker, T. Plötz, Pd disease state assessment in naturalistic environments using deep learning, in: Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015.
- [21] N.Y. Hammerla, S. Halloran, T. Ploetz, Deep, convolutional, and recurrent models for human activity recognition using wearables, IJCAI, 2016.
- [22] J. Hannink, T. Kautz, C.F. Pasluosta, K.-G. Gaßmann, et al., Sensor-based gait parameter extraction with deep convolutional neural networks, *IEEE J. Biomed. Health Inform.* 21 (1) (2017) 85–93.
- [23] T. Hayashi, M. Nishida, N. Kitaoka, K. Takeda, Daily activity recognition based on DNN using environmental sound and acceleration signals, in: Signal Processing Conference (EUSIPCO), 23rd European, 2015, pp. 2306–2310.
- [24] G.E. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets, *Neural Comput.* 18 (7) (2006) 1527–1554.



- [25] L. Hu, Y. Chen, S. Wang, J. Wang, J. Shen, X. Jiang, Z. Shen, Less annotation on personalized activity recognition using context data, in: *UIC*, 2016, pp. 327–332.
- [26] M. Inoue, S. Inoue, T. Nishida, Deep recurrent neural network for mobile human activity recognition with high throughput, (2016). arXiv preprint arXiv:1611.03607.
- [27] W. Jiang, Z. Yin, Human activity recognition using wearable sensors by deep convolutional neural networks, in: *MM*, ACM, 2015, pp. 1307–1310.
- [28] A. Khan, S. Mellor, E. Berlin, R. Thompson, R. McNaney, P. Olivier, T. Plötz, Beyond activity recognition: skill assessment from accelerometer data, in: *UbiComp*, ACM, 2015, pp. 1155–1166.
- [29] U.M. Khan, Z. Kabir, S.A. Hassan, S.H. Ahmed, A deep learning framework using passive wifi sensing for respiration monitoring, arXiv preprint (2017). arXiv:1704.05708.
- [30] Y. Kim, Y. Li, Human activity classification with transmission and reflection coefficients of on-body antennas through deep convolutional neural networks, *IEEE Trans. Antennas Propag.* 65 (5) (2017) 2764–2768.
- [31] Y. Kim, B. Toomajian, Hand gesture recognition using micro-doppler signatures with convolutional neural network, *IEEE Access* (2016).
- [32] N.D. Lane, P. Georgiev, Can deep learning revolutionize mobile sensing? in: *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications*, ACM, 2015, pp. 117–122.
- [33] N.D. Lane, P. Georgiev, L. Qendro, Deepear: robust smartphone audio sensing in unconstrained acoustic environments using deep learning, in: *UbiComp*, ACM, 2015, pp. 283–294.
- [34] O.D. Lara, M.A. Labrador, A survey on human activity recognition using wearable sensors, *IEEE Commun. Surv. Tutorials* 15 (3) (2013) 1192–1209.
- [35] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (7553) (2015) 436–444.
- [36] S.-M. Lee, S.M. Yoon, H. Cho, Human activity recognition from accelerometer data using convolutional neural network, in: *Big Data and Smart Computing (BigComp)*, IEEE International Conference on, 2017, pp. 131–134.
- [37] X. Li, Y. Zhang, M. Li, I. Marsic, J. Yang, R.S. Burd, Deep neural network for RFID based activity recognition, *Wireless of the Students, by the Students, and for the Students (S3) Workshop with MobiCom*, 2016a.
- [38] X. Li, Y. Zhang, I. Marsic, A. Sarcevic, R.S. Burd, Deep learning for RFID-based activity recognition, in: *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*, ACM, 2016b, pp. 164–175.
- [39] Y. Li, D. Shi, B. Ding, D. Liu, Unsupervised feature learning for human activity recognition using smartphone sensors, in: *Mining Intelligence and Knowledge Exploration*, Springer, 2014, pp. 99–107.
- [40] C. Liu, L. Zhang, Z. Liu, K. Liu, X. Li, Y. Liu, Lasagna: towards deep hierarchical understanding and searching over mobile sensing data, in: *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, ACM, 2016, pp. 334–347.
- [41] S. Mohammed, I. Tashev, Unsupervised deep representation learning to remove motion artifacts in free-mode body sensor networks, in: *Wearable and Implantable Body Sensor Networks (BSN)*, 2017 IEEE 14th International Conference on, IEEE, 2017, pp. 183–188.
- [42] F.J.O. Morales, D. Roggen, Deep convolutional feature transfer across mobile activity recognition domains, sensor modalities and locations, in: *Proceedings of the 2016 ACM International Symposium on Wearable Computers*, ACM, 2016, pp. 92–99.
- [43] A. Murad, J.-Y. Pyun, Deep recurrent neural networks for human activity recognition, *Sensors* 17 (11) (2017) 2556.
- [44] F.J. Ordóñez, D. Roggen, Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition, *Sensors* 16 (1) (2016) 115.
- [45] S.J. Pan, Q. Yang, A survey on transfer learning, in: *Knowledge and Data Engineering*, IEEE Transactions on, 22, 2010, pp. 1345–1359.
- [46] M. Panwar, S.R. Dyuthi, K.C. Prakash, D. Biswas, A. Acharyya, K. Maharatna, A. Gautam, G.R. Naik, Cnn based approach for activity recognition using a wrist-worn accelerometer, in: *Engineering in Medicine and Biology Society (EMBC)*, 2017 39th Annual International Conference of the IEEE, IEEE, 2017, pp. 2438–2441.
- [47] T. Plötz, N.Y. Hammerla, P. Olivier, Feature learning for activity recognition in ubiquitous computing, in: *IJCAI*, 22, 2011, p. 1729.
- [48] B. Pourbabaee, M.J. Roshtkari, K. Khorasani, Deep convolution neural networks and learning ecg features for screening paroxysmal atrial fibrillation patients, *IEEE Trans. Syst. Man Cybernetics* (2017).
- [49] D. Prelec, H.S. Seung, J. McCoy, A solution to the single-question crowd wisdom problem, *Nature* 541 (7638) (2017) 532–535.
- [50] J. Qin, L. Liu, Z. Zhang, Y. Wang, L. Shao, Compressive sequential learning for action similarity labeling, *IEEE Trans. Image Process.* 25 (2) (2016) 756–769.
- [51] V. Radu, N.D. Lane, S. Bhattacharya, C. Mascolo, M.K. Marina, F. Kawsar, Towards multimodal deep learning for activity recognition on mobile devices, in: *UbiComp*, ACM, 2016, pp. 185–188.
- [52] D. Ravi, C. Wong, B. Lo, G.-Z. Yang, Deep learning for human activity recognition: a resource efficient implementation on low-power devices, in: *Wearable and Implantable Body Sensor Networks (BSN)*, 2016 IEEE 13th International Conference on, IEEE, 2016, pp. 71–76.
- [53] D. Ravi, C. Wong, B. Lo, G.-Z. Yang, A deep learning approach to on-node sensor data analytics for mobile or wearable devices, *IEEE J. Biomed. Health Inform.* 21 (1) (2017) 56–64.
- [54] C.A. Ronao, S.-B. Cho, Deep convolutional neural networks for human activity recognition with smartphone sensors, in: *International Conference on Neural Information Processing*, Springer, 2015, pp. 46–53.
- [55] C.A. Ronao, S.-B. Cho, Evaluation of deep convolutional neural network architectures for human activity recognition with smartphone sensors, in: *Proc. of the KIISE Korea Computer Congress*, 2015, pp. 858–860.
- [56] C.A. Ronao, S.-B. Cho, Human activity recognition with smartphone sensors using deep learning neural networks, *Expert Syst. Appl.* 59 (2016) 235–244.
- [57] A. Sathyanarayana, S. Joty, L. Fernandez-Luque, F. Ofli, J. Srivastava, A. Elmagarmid, S. Taheri, T. Arora, Impact of physical activity on sleep: a deep learning based exploration, (2016). arXiv preprint arXiv:1607.07034.
- [58] J. Schmidhuber, Deep learning in neural networks: an overview, *Neural Netw.* 61 (2015) 85–117.
- [59] M.S. Singh, V. Pondenkandath, B. Zhou, P. Lukowicz, M. Liwicki, Transforming sensor data to the image domain for deep learning—an application to footprint detection, (2017). arXiv preprint arXiv:1701.01077.
- [60] R. Stewart, S. Ermon, Label-free supervision of neural networks with physics and domain knowledge, in: *AAAI*, 2017, pp. 2576–2582.
- [61] P. Vepakomma, D. De, S.K. Das, S. Bhansali, A-wristocracy: deep learning on wrist-worn sensing for recognition of user complex activities, in: *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, IEEE, 2015, pp. 1–6.
- [62] K.H. Walse, R.V. Dharaskar, V.M. Thakare, Pca based optimal ann classifiers for human activity recognition using mobile sensors data, in: *Proceedings of First International Conference on Information and Communication Technology for Intelligent Systems*, 1, Springer, 2016, pp. 429–436.
- [63] A. Wang, G. Chen, C. Shang, M. Zhang, L. Liu, Human activity recognition in a smart home environment with stacked denoising autoencoders, in: *International Conference on Web-Age Information Management*, Springer, 2016, pp. 29–40.
- [64] J. Wang, Y. Chen, S. Hao, W. Feng, Z. Shen, Balanced distribution adaptation for transfer learning, in: *The IEEE International conference on data mining (ICDM)*, 2017, pp. 1129–1134.
- [65] J. Wang, X. Zhang, Q. Gao, H. Yue, H. Wang, Device-free wireless localization and activity recognition: a deep learning approach, *IEEE Trans. Veh. Technol.* (2016).
- [66] J.B. Yang, M.N. Nguyen, P.P. San, X.L. Li, S. Krishnaswamy, Deep convolutional neural networks on multichannel time series for human activity recognition, in: *IJCAI*, Buenos Aires, Argentina, 2015, pp. 25–31.
- [67] Q. Yang, Activity recognition: Linking low-level sensors to high-level intelligence, in: *IJCAI*, 9, 2009, pp. 20–25.
- [68] S. Yao, S. Hu, Y. Zhao, A. Zhang, T. Abdelzaher, Deepsense: A unified deep learning framework for time-series mobile sensing data processing, in: *WWW*, 2017, pp. 351–360.
- [69] T. Zebin, P.J. Scully, K.B. Ozanyan, Human activity recognition with inertial sensors using a deep learning approach, in: *SENSORS*, 2016, pp. 1–3.
- [70] M. Zeng, L.T. Nguyen, B. Yu, O.J. Mengshoel, J. Zhu, P. Wu, J. Zhang, Convolutional neural networks for human activity recognition using mobile sensors, in: *Mobile Computing, Applications and Services (MobiCASE)*, 2014 6th International Conference on, IEEE, 2014, pp. 197–205.
- [71] L. Zhang, X. Wu, D. Luo, Human activity recognition with HMM–DNN model, in: *Cognitive Informatics & Cognitive Computing (ICCI\* CC)*, 2015 IEEE 14th International Conference on, IEEE, 2015, pp. 192–197.
- [72] L. Zhang, X. Wu, D. Luo, Real-time activity recognition on smartphones using deep neural networks, in: *UIC*, IEEE, 2015, pp. 1236–1242.
- [73] L. Zhang, X. Wu, D. Luo, Recognizing human activities from raw accelerometer data using deep neural networks, in: *IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, 2015, pp. 865–870.
- [74] S. Zhang, W.W. Ng, J. Zhang, C.D. Nugent, Human activity recognition using radial basis function neural network trained via a minimization of localized generalization error, in: *International Conference on Ubiquitous Computing and Ambient Intelligence*, Springer, 2017a, pp. 498–507.
- [75] Y. Zhang, X. Li, J. Zhang, et al., Car-a deep learning structure for concurrent activity recognition, in: *IPSN*, 2017b, pp. 299–300.
- [76] Y. Zheng, Q. Liu, E. Chen, Y. Ge, J.L. Zhao, Time series classification using multi-channels deep convolutional neural networks, in: *International Conference on Web-Age Information Management*, Springer, 2014, pp. 298–310.
- [77] Y. Zheng, Q. Liu, E. Chen, Y. Ge, J.L. Zhao, Exploiting multi-channels deep convolutional neural networks for multivariate time series classification, *Frontiers Comput. Sci.* 10 (1) (2016) 96–112.