Survey on wearable sensor modality centred human activity recognition in health care

Yan Wang[a,d], Shuang Cang[b], Hongnian Yu[c,d,*]

[a]School of Electric and Information Engineering, Zhongyuan University of Technology, Zhengzhou, 450007, China

[b]School of Economics and Management, Yanshan University, Qinhuangdao 066004, China.

[c]Industry 4.0 Artificial Intelligence Laboratory, Dongguan University of Technology, Dongguan 511700, China

[d]Faculty of Science and Technology, Bournemouth University, Poole, BH125BB, UK

[*]Corresponding author: yuh@bournemouth.ac.uk

Abstract

Increased life expectancy coupled with declining birth rates is leading to an aging population structure. Aging-caused changes, such as physical or cognitive decline, could affect people's quality of life, result in injuries, mental health or the lack of physical activity. Sensor-based human activity recognition (HAR) is one of the most promising assistive technologies to support older people's daily life, which has enabled enormous potential in human-centred applications. Recent surveys in HAR either only focus on the deep learning approaches or one specific sensor modality. This survey aims to provide a more comprehensive introduction for newcomers and researchers to HAR. We first introduce the state-of-art sensor modalities in HAR. We look more into the techniques involved in each step of wearable sensor modality centred HAR in terms of sensors, activities, data pre-processing, feature learning and classification, including both conventional approaches and deep learning methods. In the feature learning section, we focus on both hand-crafted features and automatically learned features using deep networks. We also present the ambient-sensor-based HAR, including camera-based systems, and the systems which combine the wearable and ambient sensors. Finally, we identify the corresponding challenges in HAR to pose research problems for further improvement in HAR.

Keywords: Human activity recognition, wearable sensors, deep learning, features, healthcare

1. Introduction

Globally, the population aged 60 or over is growing faster. The world population report predicts that life expectancy at birth will rise from 71 years in 2010-2015 to 77 years in 2045-2050 (Farah et al., 2019, United Nations, 2017). Most societies face problems to ensure that their health systems are ready to adapt to the demographic shift. Some measures, e.g., developing new systems with medical and assistive technologies for providing long-term care or creating age-friendly environments, have been exploring to maintain or improve older people's quality of life. These years have been witnessing the development of assistive technologies in promoting independent, active and healthy aging due to the advancement of sensors, wireless communication,

and machine learning techniques (Carmeli et al., 2016, Kon et al., 2017, Kuerbis et al., 2017). Among these technologies, sensor-based human activity recognition (HAR) becomes one of the most promising solutions to assist older people's daily life (Wang et al., 2018, Chernbumroong et al., 2013, Janidarmian et al., 2017, Lee et al., 2017, Tunca et al., 2014). HAR learns activities from a series of observations on the actions of subjects and the environmental conditions in real-life settings, which has been explored in human-centred applications, such as assisted living (De et al., 2017), interactive games (Terada et al., 2010), sport activity monitoring (Zhou et al., 2016), social physical interaction (Augimeri et al., 2010), factory workers monitoring (Huang et al., 2007), etc.

The early study on HAR can be traced back to the work by Abowd et al., 1998. Researchers initially focus on activity recognition from videos and images, but later when everyday life is considered, they start to explore tracking human behaviour by using wearable and ambient sensors (Bulling et al., 2014, Ke et al., 2013, Zolfaghari et al., 2016) as well. The progress made in HAR during the past few decades motivates researchers to improve the recognition performance and practicality of HAR under more realistic settings in different ways. HAR process is complex, roughly follows the five steps: 1), selecting and deploying appropriate sensors to a human body or the environment to capture the user's behaviour or the change of the environment where the user is performing activities; 2), collecting and pre-processing the data from the deployed sensors based on a specific task; 3), extracting useful features from the sensor data for later classification; 4), training the classification models with appropriate machine learning algorithms to infer activities; 5), testing the learning models to give decisions and performance reports. Each step above involves plenty of technologies and methods available to use and also has the corresponding research questions to tackle (Lara et al., 2013, Cornacchia et al., 2017, Nweke et al., 2018). The technologies involved in HAR can cover sensing technologies, wireless networks communicating, data pre-processing, feature learning, feature dimensionality reduction, classification or regression techniques, etc.

In terms of the sensors deployed in HAR, the existing HAR systems can be broadly categorized into three modalities: the ambient sensor-based HAR (ASHAR), the wearable sensor-based HAR (WSHAR), and the hybrid sensory-based HAR (HSHAR). ASHAR systems infer human activities from the sensors that are fixed in the environment or attached to some specific objects, such as wall, door, kettle, floor, etc., and the ambient sensors can include light sensor, reed switch sensor, Radio Frequency Identification (RFID), passive infrared (PIR), temperature, flow sensor, pressure sensor, (Zhang et al., 2017, Debes et al., 2016, Mehr et al., 2016, Tunca et

al., 2014, Liu et al., 2018), etc. ASHAR sensor modality is less obtrusive because of no on-body sensors deployed, while usually at the price of poor flexibility and complex sensor deployment in homes. ASHAR works in a limited area where the sensors are deployed. Besides, systems using pure normal ambient sensors may fail to function in some situations when the user does not contact the objects attached with ambient sensors or does not enter the functioning area of a sensor installed in the environment.

The alternative to ASHAR with fixed sensor deployment is WSHAR, which identifies human activities by mining the informative data from wearable sensors using machine learning algorithms. WSHAR can function in a relatively large space when the wearer is moving. Currently, smartphones, smartwatches, smart clothes, and other specifically-designed devices are the mainstream products embedded wearable technologies in HAR (Hassan et al., 2018, Filippoupolitis et al., 2017, Adaskevicius, 2014). Generally, placing more sensors on multiple body parts (e.g., head, wrists, waist, legs, feet) can benefit improving the performance and robustness of WSHAR (Laudanski et al., 2015, Gao et al., 2014, Chernbumroong et al., 2014). However, multiple sensors with complex sensor deployment on body could cause higher costs, practical deployment difficulties, and obtrusions for older users especially those who can live independently. Meanwhile, pure WSHAR systems also have some limitations that may enable less accurate recognition for certain activities that contain similar sensor-derived attributes, such as brushing and eating (Chernbumroong et al., 2013).

ASHAR and WSHAR have their own strengths and weaknesses. It is shown that combining different sensor modalities can improve recognition accuracy (Cornacchia et al., 2017). For example, Logan et al., 2007, StikicVan Laerhoven et al., 2008 present the improved activity recognition performance by combining the wearable sensors with the infrared sensors. Roy et al., 2016 use ambient and mobile data in a multi-inhabitant environment for daily activities detecting. The initial results can reach around 70%, which is much higher than the results by using the smartphone-based accelerometers alone. It is obvious that the combination of sensor modalities can capture rich information about human activities, thereby improving the performance of HAR. Nevertheless, HSHAR could increase the cost and complexity of a HAR system compared with a single sensor modality. HSHAR also needs data fusion and sensing synchronization from different sensor modalities. Among the three sensor modalities, WSHAR attract more attention due to its low cost, flexibility in daily use and satisfied performance (Roy et al., 2016, Diethe et al., 2017), and has enabled enormous applications in assisted living, such as gait analysis (Anwary et al., 2018), rehabilitation (Hermanis et al., 2016), fall detection (Jung et

al., 2015), sports assessment (Um et al., 2016), daily activity analysis (Y Wang et al., 2018), etc. This survey then focuses on WSHAR and also looks at ASHAR and HSHAR.

The state-of-art surveys in HAR are either focusing on the deep learning approaches (Wang et al., 2017, Nweke et al., 2018) or only each single sensor modality (Cornacchia et al., 2017, Morales et al., 2017). This survey focuses on the wearable sensor-based HAR and keeps an eye on other sensor modalities. Specifically, we detail the techniques involved in each step of wearable sensor-based HAR in terms of sensors, activities, data pre-processing, feature learning and classification. Both the hand-crafted and automatically learned features are investigated in the feature learning section. The survey can provide strong clues for new researchers who might be in a dilemma about system designing or methods choosing in HAR and fills the gaps of no comprehensive surveys which include both conventional and deep learning methods in HAR. The survey pipeline is shown in Fig.1.

The rest of the paper is organised as follows. Section 2 focuses on the wearable sensor-based HAR. Section 2 has seven subsections and in each subsection, the descriptions, strengths, and limitations of the reviewed approaches are discussed. Section 3 surveys the ambient sensor-based HAR including camera-based HAR and the hybrid sensory HAR which combine two or three sensor modalities. Section 4 presents the performance evaluation and applications of HAR in health care. Section 5 concludes the survey and poses some research challenges in HAR for further research.

2 Wearable sensor-based HAR (WSHAR)

2.1 Overview of WSHAR

Development of wearable devices, such as smartwatches, smartphones, wristbands, smart clothes, makes it feasible to acquire data from the ubiquitous equipment and provide continuous monitoring of human activities (Adaskevicius, 2014, Filippoupolitis et al., 2017, Hassan et al., 2018). Data-driven-based WSHAR systems basically share a similar procedure, as shown in Fig.2. Flowchart A in Fig.2 presents the process using conventional approaches to realize HAR, in which the features are generated manually according to expert knowledge (Chernbumroong et al., 2014, Sani et al., 2017). First, the raw data from multiple types of body-worn sensors (accelerometer, gyroscope, heart rate sensor, etc.) are obtained at a certain sampling rate and then transmitted to a processing centre (laptop, tablet, smartphone, etc.) through specific communication
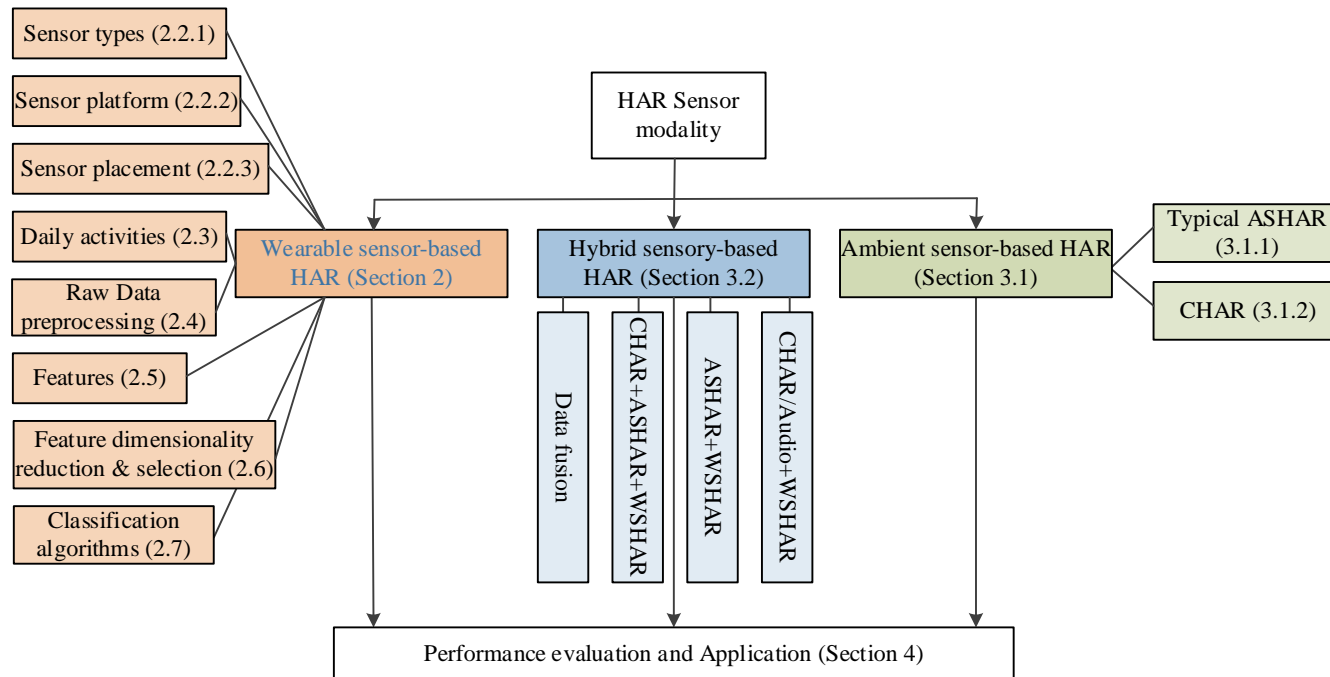
Fig.1 Survey pipeline

(CHAR-Camera-based HAR, ASHAR-Ambient sensor-based HAR, WSHAR-Wearable sensor-based HAR, HSHAR-Hybrid sensory-based HAR, 2.2.1 is subsection 2.2.1)
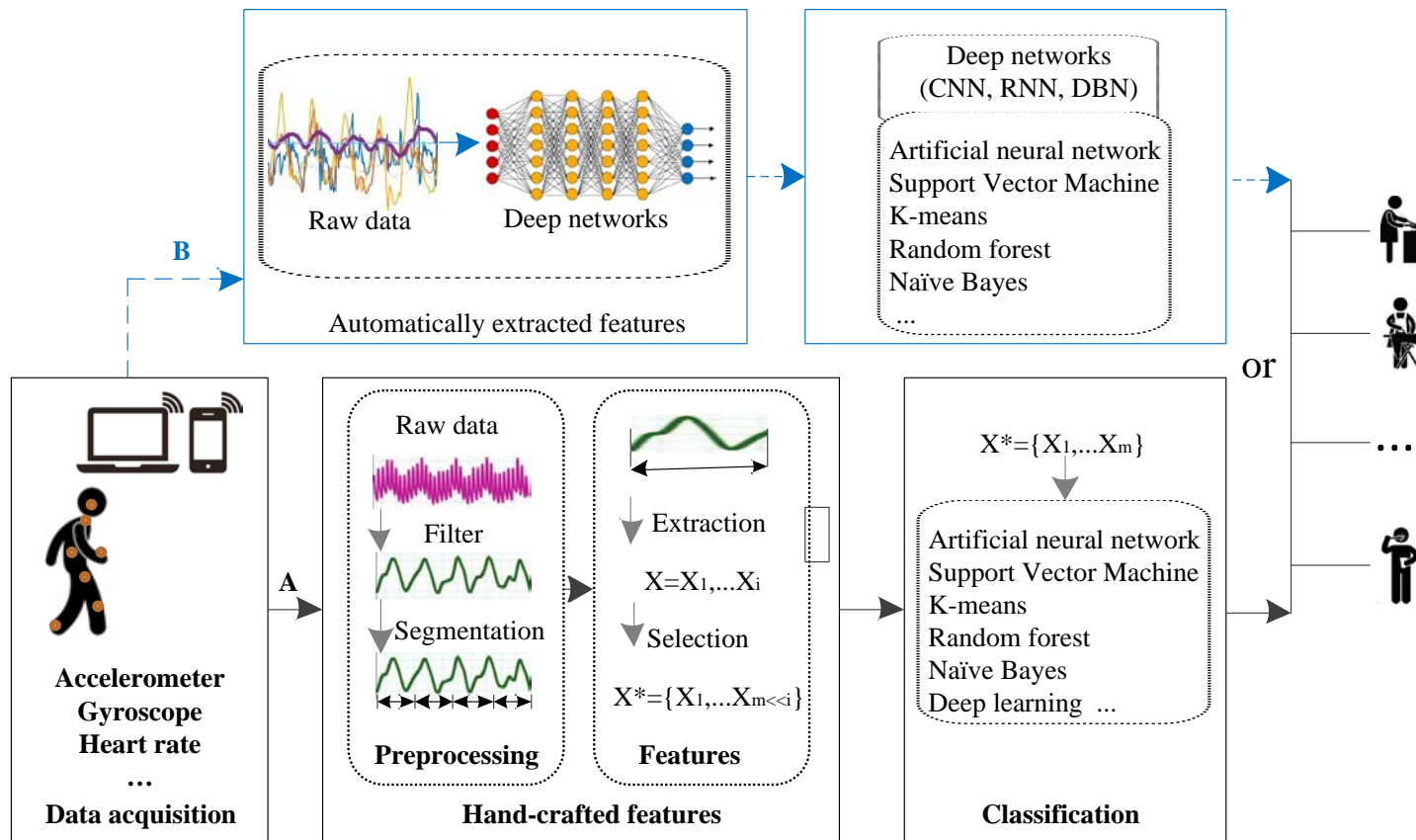
Fig.2 Learning procedure of WSHAR

(RNN-Recurrent Neural Network, CNN-Convolutional Neural Network, RBM-Restricted Boltzmann machine)

technologies (Bluetooth, Zigbee, Wi-Fi, etc.); the pre-processing stage mainly involves filtering and segmenting the raw data; then informative features are extracted in a hand-crafted way (mean, variance, dominant frequency, entropy etc.); followed by applying the specific feature dimension reduction techniques or feature selection algorithms to obtain the optimal and smaller-size feature set for further learning and computation burden reducing; finally, the optimal feature set is fed to the classifiers for classification models training and testing. Flowchart B in Fig.2 instead gives the typical process of using deep leaning methods for HAR, in which the features can be learned automatically from different types of deep networks, such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deep Belief Network (DBN), Restricted Boltzmann machine (RBM) (Plötz et al., 2011, Panwar et al., 2017). The feature learning and learning model building in flowchart B are often performed simultaneously with these deep networks.

2.2 Wearable sensors

2.2.1 Sensor type

The advances in sensors make it possible and feasible to explore assisted living in health care and wellbeing with wearable sensors. Wearable sensors, different from the common-used industrial sensors, are designed to meet some specific requirements: high integration density, small size, low power consumption as well as high measurement accuracy, etc. The sensors are integrated into a small-size device for being conveniently attached to the user's body parts. Wearable sensors can include inertial sensors, physical health sensors, environmental sensors, camera, microphone, etc. Table 1 presents the most popularly used wearable sensors in HAR. Among them, motion-based inertial sensors have been well applied in WSHAR, such as accelerometer, gyroscope or magnetometer, which are capable of detecting and measuring acceleration, angular velocity, magnetic fields, tilt, shock, vibration, rotation, and multiple degrees-of-freedom motion (Chernbumroong et al., 2014, Gjoreski et al., 2011b, Hassan et al., 2018). These observations vary sensitively along with a wearer's movement or body postures, thereby delivering rich motion-caused information. Kwapisz et al., 2011 utilize accelerometers to identify five physical activities, i.e., walking, jogging, ascending/descending stairs, sitting and standing. Deng et al., 2014 develop a fast and robust activity recognition model based on Reduced Kernel Extreme Learning. Guo et al., 2016 use an accelerometer, a magnetometer, and a gyroscope built in a smartphone for patients' activity recognition. Inertial sensors still suffer from some limitations, e.g., the calibration for effective measurements, battery life limitation due to continued logging, or arbitrary signals associated with activity performing.

Table 1 Wearable sensors used in HAR

Wearable sensors	Examples	Pros	Cons
Inertial sensors	Accelerometer (Chernbumroong et al., 2014, Hassan et al., 2018) Gyroscope (Anwary et al., 2018) Magnetometer (Gjoreski et al., 2011b)	Well applied, delivering rich motion information, small size, easy to use, etc.	Battery life limitation, arbitrary signals companied with activities, etc.
Physical health sensors	Electrocardiogram (ECG) (Zhang et al., 2018) Skin temperature (Yoon et al., 2016) Heart rate (HR) (Tapia et al., 2007, Mehrang et al., 2017) Electroencephalograph (EEG) (Nakamura et al., 2010) Electromyogram (EMG) (Georgi et al., 2015) Force/pressor sensor (Lorussi et al., 2016)	Delivering rich vital signals related to activities, can be used for rehabilitation and health condition detection, etc.	Unable to obtain large-scale application due to the issues of size, precision, price, etc.
Environmental sensors	Temperature (Chernbumroong et al., 2014) Humidity (Parkka et al., 2006) Light sensor (Bhattacharya et al., 2016) Barometer, etc. (Wang et al., 2018)	Delivering context information related to activities	Usually used with inertial sensors and producing noise signals, etc.
Others	Camera (Zhan et al., 2012) Microphone (Fontana et al., 2015) GPS, etc. (Reddy et al., 2010)	Complementary information with other sensors	Privacy concerns, complex algorithms applied, etc.

ECG, blood glucose (BG), respiratory rate (RR), etc., are used sometimes with the inertial sensors to recognize the activities with rehabilitation purpose or capture vital signals for health condition evaluation. Chen et al., 2014 develop a framework to detect epileptic seizures using EEG sensors. Chernbumroong et al., 2014 propose a practical activity recognition system by combining a heart rate sensor attached to the chest with another six sensors worn on the wrists. Physical sensors have not been unable to obtain a large-scale application in WSHAR due to the problems of size, precision, price, etc.

For environmental sensors, only temperature sensors, barometers as well as light sensors can be often found in HAR. For example, Maurer et al., 2006 implement a multi-sensor platform embedded with a light sensor. They attach the platform on five different positions to explore the best location on body achieving the highest accuracy. A smartphone-based barometer is used to help detect a total of 15 activities with other sensors inside (Khan et al., 2014).

2.2.2 Sensor platform

In WSHAR, the sensors are typically integrated into one platform carried by users when they perform activities. To minimize the obtrusiveness during use, the sensor devices are often seen in the following modes: smartphones, smartwatches, smart clothes, inertial units, specifically-designed platforms, etc.

Today's smartphones are well equipped with a variety of sensors (such as accelerometers and gyroscopes) and are ubiquitously carried by people everywhere and every day. Using the data acquired from these sensors could enable applications to recognize a wide range of daily activities (Hassan et al., 2018, Kwon et al., 2014, Guo et al., 2016, Reddy et al., 2010, Sun et al., 2010). Also, smartphones are equipped with memory and battery, which provides a system for HAR without additional hardware requirements. The main problems when using smartphones for HAR involve the constraints of limited sensor types and locations (pockets, belts or bags). Meanwhile, carrying a smartphone on body all the time might not be suitable for everyday use when the phone carrier performs daily activities at home. Furthermore, retraining procedures or transforms of coordinate are normally needed to achieve HAR due to arbitrary orientations of the way of smartphone carrying (Sun et al., 2010, Morales et al., 2014).

Smartwatches are designed with integrated sensors that enable a connection to a PC or a phone. The typical examples of using smartwatches to identify daily activities could be seen in Filippoupolitis et al., 2017, Vepakomma et al., 2015, Chernbumroong et al., 2014, Mortazavi et al., 2014 and so on. A smartwatch is typically wrist-mounted. With a relatively standard and fixed body location, wearing a smartwatch is more
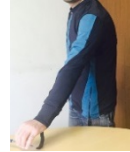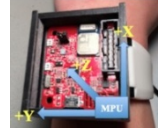
convenient and less obtrusive for the user compared to carrying a smartphone all the time. Nevertheless, smartphones and smartwatches share the same problem that the sensors inside them are fixed and might not be the exact ones required for a specific task. In some cases, the data from them might not be open-source.

Smart clothes can embed more sensors, especially physical sensors, to achieve more information. They are often found in long term monitoring applications due to the easy wearing (Adaskevicius, 2014). For instance, Smart shirts are designed to monitor precise cardiac, respiratory, sleep and other daily activities, which incorporate heart rate and ECG sensors (Hexoshin, 2018). Lorussi et al., 2016 develop a smart textile platform, including sensing shirt, sensing trousers, sensing gloves and sensing shoes for the assessment of stroke patients. The platform embeds or knits inertial sensors, textile goniometers, piezoresistive sensors, EMG and goniometers. Zhou et al., 2016 use two types of textile-based sensors: a fabric pH sensor to collect and analyse sweat and piezoresistive textiles to capture body movements. Smart clothes are also designed to track babies' sleep, breathing, body position (Mimobaby, 2018). The abovementioned smart clothes are usually needed to wear tightly to ensure the quality contact of the sensors with the skin or the body parts, which may affect the comfort of the wearer for daily use. On the other hand, the relative movement between the body parts and the sensors due to the loose wearing will give rise to motion artefacts.

An inertial measurement unit (IMU) is a special device that measures and reports a craft's velocity and orientation, using a combination of an accelerometer, a gyroscope, a magnetometer and sometimes together with a barometer. One or some combinations of the IMU sensors are often employed to detect human gestures or activities in different applications and show the satisfied performances (Georgi et al., 2015, Montalto et al., 2015, Bulling et al., 2014, Su et al., 2014).

Specifically-designed platforms are built for specific or common research purposes in HAR, in which the sensors required for a specific task are integrated. Burns et al., 2010 design a flexible sensing device with multiple sensors. Their device contains the capabilities of kinematic sensing, physiological sensing, ambient sensing and external hardware integration. Uddin et al., 2015 present a framework with a wrist-worn 9-axis-sensor device. They verify the feasibility of the device based on hands washing and drinking. Cook et al., 2015 design an open-source, wearable, eight-channel bio-potential data collection platform integrated with an ECG and an accelerometer sensor, which can be used to record health-related information. Specifically developed sensor devices can meet the sensor requirements for a specific task, while it may mean an extra cost in hardware and research period. The popular sensor platforms used in WSHAR are summarized in Table 2.

Table 2 Sensor platforms in WSHAR

Platform	Case studies	Strengths	Weaknesses	Picture
Smartphones	Sun et al., 2010 Guo et al., 2016 Hassan et al., 2018	Ubiquitous, equipped with a variety of sensors, battery and memory	Limited placing locations on body, arbitrary orientations in pockets, etc.	
Smartwatches	Vepakomma et al., 2015 Chernbumroong et al., 2014 Uslu et al., 2013	Integrated sensors, a relatively standard and fixed body location	Limited sensor types for different applications	
Smart clothes	Adaskevicius, 2014 Hexoshin, 2018 Lorussi et al., 2016	More sensors embedded, long term monitoring, the relative movement between the body parts and the sensors, etc.	Usually needed to wear tightly to ensure the quality contact of the sensors with the skin or other body parts	
Inertial measurement unit (IMU)	Georgi et al., 2015 Su et al., 2014 Anwary et al., 2017	A fixed combination of sensors, small, low power, can also provide the attitude angles of the device, etc.	Time-consuming alignment and calibration, etc.	
Specifically-designed devices	Y Wang et al., 2018 Uddin et al., 2015 Cook et al., 2015	The sensors exactly required for a specific task or a common research purpose in HAR	An extra cost in hardware and research period	

2.2.3 Sensor placement

Sensor placement refers to the body locations where the sensors are placed and how the sensors are attached to those locations, which is a research-worthy problem in WSHAR. Sensor placement may vary along different applications. For example, a foot-mounted accelerometer can well reflect the foot or leg involved motion, thereby for gait, step, distance or energy consumption detection (Anwary et al., 2018, Chamroukhi et al., 2013, Moncada-Torres et al., 2014, Vepakomma et al., 2015). The wrist-worn sensors can help recognise normal activities, such as ironing, brushing teeth and cooking (Mannini et al., 2010, Chernbumroong et al., 2013). The thigh-located sensors are sensitive to the leg-involved activities, like jogging, riding, walking, running, etc.(Wu et al., 2012, Moncada-Torres et al., 2014, Ronao et al., 2015). Most potential body locations are explored to place sensor(s): hand (Kundu et al., 2017), arm (Bulling et al., 2014), wrist (Pavey et al., 2017), chest (Gao et al., 2014), pocket (KwonKwon et al., 2014), head (He et al., 2014), feet (Anwary et al., 2018), shank (Bahrepour et al., 2011), thigh (Banos et al., 2013), trunk (Bahrepour et al., 2011), vest (Bourke et al., 2008), waist (Barreto et al., 2014), ankle (Suto et al., 2017), belt (Capela et al., 2015), pelvic (Ravi et al., 2005), hip (Banos et al., 2013), leg (Wang et al., 2013), abdomen (Zheng et al., 2013), back (He et al., 2014), knee (Atallah et al., 2010), ear (Pansiot et al., 2007), neck (Fontana et al., 2015), etc.

In terms of the sensor placement, we categorize WSHAR into four cases: the first places one single sensor on one single body part (**One to One**). One to One sensor placement aims to build a basic wearable framework for HAR. In this case, the sensor's location may vary with tasks, from the head to the feet, but fixes on one body part. Suto et al., 2017 investigate the efficiency of the popular machine learning strategies based on a right-ankle-mounted accelerometer, and their results suggest that one sensor is not enough for appropriate daily activity recognition due to the similar data generated from one sensor for different activities. The second case attaches one single type of sensor on multiple body parts to gain complementary information from different body parts (**One to Multi**). One to One sensor placement might deliver limited information for HAR; researchers then place the accelerometers to multiple body parts with the aim of capturing richer information or evaluating the contributions of different sensor positions to recognition performance. Sztyler et al., 2017 develop a position-aware HAR system by placing seven accelerometers in different body positions. The third case places a sensor device with two or more type of sensors built-in on only one body part (**Multi to One**), with the aim of capturing diverse-source information from different sensors compared to One to One case. Vepakomma et al., 2015 propose a novel framework for human activity recognition. They use a wrist-worn device with multiple sensors

inside, including accelerometer, gyroscope, barometric pressure, humidity, etc. These multi-modal sensor data from the wrist-worn sensors provide rich information for recognizing complex in-home activities. The fourth case deploys multiple devices, each embedded with two or more types of sensors, on multiple body parts (**Multi to Multi**) to take the advantages of the first three cases above, which is expected to be the most comprehensive structure to achieve higher performance in WSHAR. Chernbumroong et al., 2014 present a practical home-based HAR which use multiple types of sensors on multiple body positions. They exploit seven sensors (i.e., the altimeter, accelerometer, heart rate monitor, barometer, gyroscope, light and the temperature sensor) towards activity classification.

WSHAR systems deploy a wide variety of sensors on different body parts targeting specific aims and applications. Generally, One to One is the basic deployment and more suitable for the basic recognition tasks, such as step counting or sleep quality monitoring. Placing more sensors on multiple body parts is intuitively beneficial for improving the performance and robustness, whereas this can also result in increased complexity in deployment and computation cost. Also, the sensors spread over a human body hinder the wearer doing everyday activities, this may cause the user to reject to wear them. Consequently, exploring the way to implement WSHAR with less obtrusiveness, affordable cost as well as higher accuracy becomes more significant.

2.3 Activities of daily living

HAR is an extensive research field of machine learning. Most studies in HAR focus on indoor activities of daily life (ADL) in assisted living applications (Anwary et al., 2017, Hannink et al., 2017, Jung et al., 2015). The activities in HAR can be generally grouped in three levels according to their duration and complexity: transition activities, basic activities, and complex activities. Transition activities are the temporal patterns among activities, such as stand-to-sit, sit-to-lie, push-ups, bicep curls and so on (Reyes-Ortiz et al., 2016, Mortazavi et al., 2014). The recognition of transition activities is commonly seen in fitness or rehabilitation-related applications (Masse et al., 2016, Farah et al., 2019), additionally, which can also be used to recognize complex or basic activities as a mid-level features for later classification (Y Liu et al., 2016). Basic activities are the activities which have a longer duration than transition actions, such as walking, running, lying, cooking, stairs using, etc. (Lorussi et al., 2016, Y Wang et al., 2018, Hassan et al., 2018). Complex activities are in the form of sequential, interweaved or concurrent patterns of transition or basic activities, such as coffee time, relaxing, smoking, talking and so on (Y Liu et al., 2016, Shoaib et al., 2016).

The different levels of activities and daily routine can help reveal people's daily context and safety conditions. The recognition of ADL is expected to understand, maintain and assist the daily life of the observed. For example, long-term sedentary activities may imply one person is suffering certain cognition problems or having early dementia symptoms; more sleep at daytime or less at night may reflect insomnia or other medical and psychiatric problems; frequent use of the toilet or frequent drinking are probably associated with diabetes or kidney diseases. And changes in routines prompt us that certain disorder may be happening compared with the normal patterns; on the other hand, regular eating, regular exercise, and other well-organized daily activities can reveal the subject is leading a healthy lifestyle. Also, older people living alone have a high risk of possible falls, which is the main concern for both themselves and their families. The more details about fall detection in WSHAR can refer to S S Khan et al., 2017 and Pang et al., 2019. The above-described ADLs and conditions all can be detected by HAR systems and the corresponding decisions made by the systems can be provided to assist older people living independently. Table 3 presents some case studies regarding different activity levels based on the defined activities in HAR applications.

Real-world data is the first material and crucial for the recognition tasks after determining sensor types and sensor deployment. While data acquisition can be tedious and cumbersome work, researchers may face a series of problems when collecting real-world data, such as the obtrusiveness, the ease of using sensors, the time arrangement, the experiment environment, the cost, the annotation, etc. The real-world data for a specific task should involve as more as possible target population with diverse age, gender, weight, height and health conditions. Whilst, due to the time cost and the subjects' will, the number of recruited volunteers for data collection are usually highly limited, for example, 1 in Alvarez-Alvarez et al., 2013, 12 in Bhattacharya et al., 2016, 30 in Fontana et al., 2015, 45 in Hajihashemi et al., 2013, apart from some benchmark datasets with larger population. As for the older participants, the number of participants is smaller (Bergmann et al., 2012, Chernbumroong et al., 2013, Y Wang et al., 2018).

The protocol of data collection also affects the recognition performance, and the factors can involve the number of activities, the number of participants, performing activities in a natural way or a constrained way, a controlled environment or a real home setting, etc. Some studies collect their data based on the predefined activities under a controlled environment. E.g., Laudanski et al., 2015 ask the volunteers to perform the same activity in the similar frequency and intensity, thereby achieving high performance due to the high intra-class settings. With respect to data annotation, most studies supervise the data collection process, label the data by

Table 3 Case studies in terms of activity types in WSHAR

Activity level	Application	Defined activities	Reference
Transition	Fitness	Bicep curls, crunches, push ups, jumping jacks, shoulder lateral raises	Mortazavi et al., 2014
	Rehabilitation	Loading response, push-off, swing, terminal swing	Farah et al., 2019
	Fitness	Hammer-curl with dumbbell, push-ups, etc.	Um et al., 2016
	Gait analysis	Gait	Hannink et al., 2017
	Dietary intake	Bite, drink, utensiling, etc.	Ramos-Garcia et al., 2013
	Physiatric rehabilitation	Joint dynamics, posture, head position	Hermanis et al., 2016
Basic	ADL	Brush, exercise, eat, iron, read, lie, wipe, falls, watch TV, etc.	Wang et al., 2018
	ADL and Falls	Walking, sitting, falls.	Rasheed et al., 2015
	ADL and heart failure	Standing, walking, ascending/descending stairs, heart failure, etc.	Zheng et al., 2014
	Assessment of stroke patients	Handshake, shoulder touch, etc.	Yu et al., 2016
	Fall detection	Walking, sit down, stepping up/down, running, falling, etc.	Jung et al., 2015
	ADL	Sitting, walking, stand-to-sit, sit-to-lie, etc.	Hassan et al., 2018
Complex	ADL	Relaxing, coffee time, early morning, clean up, sandwich time	Liu et al., 2016
	ADL	Walk, jog, bike, write, coffee, smoke, eat, etc.	Shoaib et al., 2016
	ADL and fitness	Sit, walk, row, jump, cycling, exercise, coffee time, etc.	Liu et al., 2016

by observers or record the process with a camera to avoid mislabelling (Deng et al., 2014). To provide a more natural environment for participants and minimize the burden of annotation, Adaskevicius, 2014 utilize a semi-automatic approach for data collection.

Researchers collect the data for their specific research purposes. They also can use the public datasets available for HAR to evaluate their proposed methods or compare their methods with other studies on the same datasets. The commonly used datasets in WSHAR are as follows, 1): *PAMAP2* (Reiss et al., 2012), which comprises daily activities (sitting, watching TV, jogging, etc.) collected from 9 elderly subjects with three inertial sensors and heart rate placed on ankle, chest, and dominant arm; 2): *SBHAR* (Anguita et al., 2013), which is originally created for six different human activities using a waist-mounted smartphone from 30 subjects and is updated to include six more postural transitions (Reyes-Ortiz et al., 2016); 3): *mHealth* (BanosGarcia et al., 2014), which covers 12 daily activities for health monitoring using three inertial sensors and ECG sensor; 4):

WISDM (Kwapisz et al., 2011), which is a dataset collected from 29 users with single accelerometer embedded in a mobile phone, including sitting, jogging, standing, working, etc.; 5): *REALDISP* (Baños et al., 2012), which is produced in gradual sensor displacement conditions, including 33 fitness activities recorded by nine wearable IMUs on different body parts from 17 subjects; 6): *MobiAct* (Vavoulas et al., 2016), which comprises data of nine different types of ADLs from 50 subjects and four different types of falls from 44 subjects using the smartphone-based accelerometer, gyroscope and orientation sensors located in a trousers' pocket; 7): *OPPORTUNITY* (ChavarriagaSagha et al., 2013), which comprises a set of basic and complex activities collected from four subjects in an environment with both ambient and wearable sensors; other benchmark datasets can refer to the survey by Wang et al., 2018.

2.4 Raw data pre-processing

The preprocessing of the collected data in Fig.2 can include filtering (noise elimination), nominalization, and segmentation, etc. This section only talks about data filtering and segmentation.

2.4.1 Filtering

In HAR, filtering is applied to the raw sensor signals to remove some unwanted components from a signal, since raw sensor data might be contaminated by electronic noise or other artefacts. Filtering is normally performed before the time series are split into time windows for feature extraction. Kalantarian et al., 2015 and Nam et al., 2013 use the low-pass filter to smooth or remove the outliers. Machado et al., 2015 apply a second-order Butterworth High-Pass filter with cut-off frequency of 0.25 Hz to isolate the body acceleration component. Hu et al., 2014 exploit the median filter for data pre-processing. N-point moving-average filters are adopted by Adaskevicius, 2014. Hassan et al., 2018 apply the median and low-pass Butterworth filter to remove the noise from the acceleration signal. On the other hand, filtering is not always applied since some researchers state that filtering may cause the loss of useful information (Atallah et al., 2007, Ordóñez et al., 2013, Fontana et al., 2015).

2.4.2 Window Segmentation

The time series data from wearable sensors are in the order of seconds or minutes which is a relatively long period compared with the sensors' sampling rate (mostly varying from 20Hz to 100Hz). For facilitating the later learning, time series are often segmented into certain time windows. The sliding window is the most popular segmentation approach due to its implementation simplicity. Sliding windows partition the time series into fixed-size windows.

Different window sizes are employed in WSHAR, which are found to vary from 0.08s (Berchtold et al., 2010), 0.1s (Murao et al., 2014), 0.2s (Zhang et al., 2012), 0.5s (ChavarriagaBayati et al., 2013), 1s (Bulling et al., 2014), 1.6s (Suto et al., 2016), 2s (Laudanski et al., 2015), 2.56s (Hassan et al., 2018), 3.88s (Chernbumroong et al., 2014), 4s (Wang et al., 2013, 5s (Machado et al., 2015), 6.7s (Bao et al., 2004), 8.53s (Guo et al., 2012), 9s (Kalantarian et al., 2015), 10s (Catal et al., 2015), 12.8s (Wang et al., 2018) to 30s (Liu et al., 2012) and even bigger. Usually, a window covers several seconds time interval. A small-size window allows for faster feature extraction in later steps but may not cover enough circles of one activity. A large-size window can cover more circles of one activity and contain the information from more than one activity; this may delay recognition. Some researchers determine the window size by using empirical values or referring to other similar studies; others try a range of lengths on their data to find the optimal size. Finding the optimal window size is a case-based task. Hu et al. 2014 conclude that the length of the window should satisfy two conditions: 1) at least one cycle of the activities is statistically included in one window and it is proved that a window of several seconds can sufficiently capture circles of activities such as walking, running, using stairs, etc.; 2) the size should be set to 2^n thereby being easily employed in the Fast Fourier Transform (FFT) algorithm in one window. Therefore, some studies which use frequency-domain features set the samples in one window as 2^n in each segment (Guo et al., 2012, Bayat et al., 2014, Wang et al., 2018).

We need to consider the sampling rate of sensors when talking about the number of samples in one window since the sample number is determined by both the window size and the sampling rate. A wide range of sampling rates are explored in WSHAR, varying from 1hz (Zhang et al., 2014), 5hz (Alshurafa et al., 2014), 6hz (Gjoreski et al., 2011b), 10hz (Nam et al., 2013), 20hz (Wang et al., 2018, Suto et al., 2016), 33hz (Chernbumroong et al., 2014), 50hz (Biswas et al., 2015, Hassan et al., 2018), 64Hz (Hammerla et al., 2016), 100hz (Sani et al., 2017), 120hz (Laudanski et al., 2015), 126hz (Gupta et al., 2014), 135hz (Dalton et al., 2013), 200hz (Yao et al., 2017), 256hz (Chen et al., 2014), and up to 800hz (Montalto et al., 2015). Generally, higher sampling rates can catch more information details but coupled with higher energy requirements and higher noise impact; lower sampling rates save considerable energy but might omit certain relevant information, thus lower accuracy. Gao et al., 2014 find based on their experimental results that the wearable systems adopting multiple sensors are less sensitive to the sampling rate than those only using a single sensor. Although the high sampling rate may help increase the recognition accuracy, it also leads to a several-fold increase in computing load. Therefore, they suggest 20 Hz to be the appropriate sampling rate for the wearable system using multiple sensors.

The number of the samples in one window versus the window size based on the reviewed works is plotted in Fig.3, with several less commonly-used numbers being excluded (Machado et al., 2015). We can see two obvious trends from Fig.3: one is that most sample numbers in one window fall into between 32 (Suto et al., 2016) and 256 (Hu et al., 2014); the other is that sample numbers of the nth power of 2 are often applied, such as 64 (Murao et al., 2014), and 128 (Ronao et al., 2016). The sampling rate as well as the trade-off between recognition efficiency and performance should be considered when manually determining the window size.

When applying window segmentations, the overlap between two consecutive windows is usually adopted to reduce information loss at the edges of the window. The most commonly used overlap rate is 50% (Laudanski et al., 2015, Kwon et al., 2014, Davis et al., 2016). There are some other studies without performing an overlap between windows (Davis et al., 2016, Banos et al., 2012).
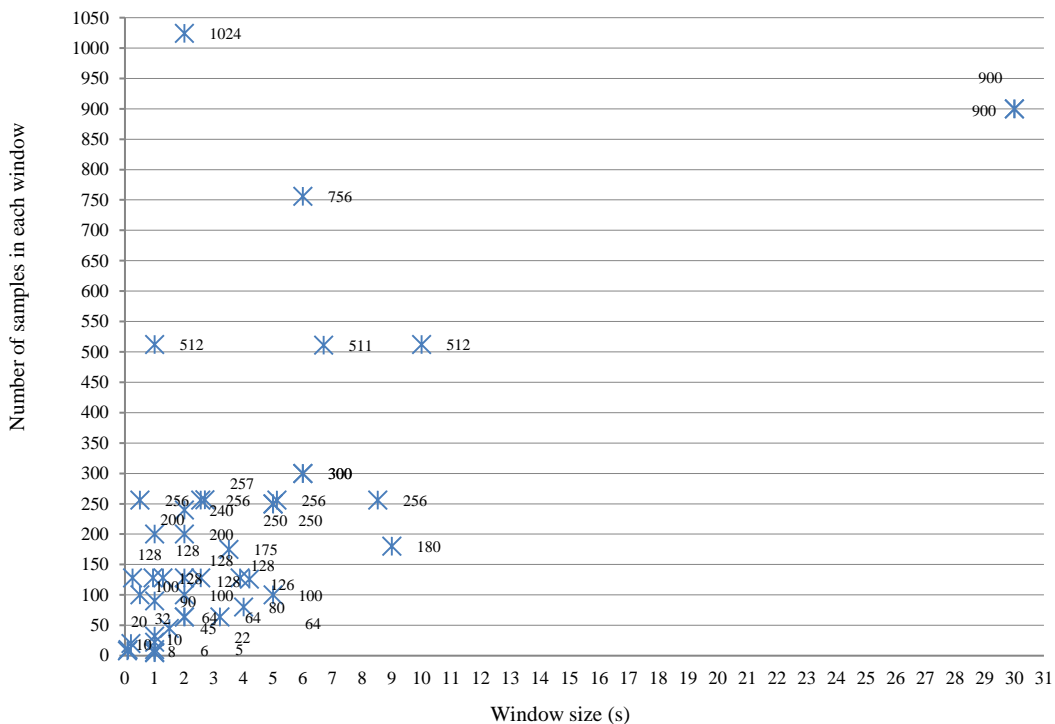
Fig.3 Sample number in each window versus window size

2.5 Features for classification

Features are the inputs for most machine learning classifiers. In general, there are two ways to extract features from raw sensor data, one is handcrafting features based on domain knowledge (Vepakomma et al., 2015) and the other is automatically learning features by deep networks (Ronao et al., 2016). Hand-crafted features are the measures computed from each window segmentation in a time domain or frequency domain, which are designed to capture the useful representation of the data for distinguishing different activities in HAR, such as mean,

median and principal frequency (Hassan et al., 2018, Suto et al., 2016). Hand-crafted features have achieved great success in HAR applications (Li et al., 2009, Hassan et al., 2018). The key advantage of using hand-crafted features is that the features are computationally lightweight to implement especially in ubiquitous devices (Morales et al., 2017). These years, deep learning approaches have been applying in HAR to automatically learn features for HAR (Hammerla et al., 2015, Sani et al., 2017). The strengths of the automatically learned features by the deep networks are that the learning can be very deep, and the learning process does not rely on domain knowledge.

2.5.1 Hand-crafted features

In the raw data space, the specific value at a specific time instant of a sample (e.g. the reading of 30℃ from a temperature sensor) does not carry sufficient information to describe an activity that the reading originates from. Furthermore, when we compare two activities in terms of two given time windows, it is nearly impossible that two time series (i.e., segmented windows) contain identical signals even the two windows represent the same activity performed by the same person. Accordingly, quantitative and informative variables can be calculated based on each window from raw sensor data; these are hand-crafted features. Consequently, hand-crafted features are elaborately designed for comparing and differentiating different activities. A wide range of hand-crafted features have been explored to improve HAR performance (Wu et al., 2012, Attal et al., 2015, Wang et al., 2016, Sani et al., 2017, Wang et al., 2018). We categorize the hand-crafted features as the following types, i.e., time-domain features, frequency-domain features, and other hybrid features.

Time-domain features are those features obtained directly from a window of sensor data and are typically statistical measures. They have been intensively investigated in different applications and proved to be effective and useful for HAR. These features are based on a comprehensive and intuitive understanding of how a specific activity or posture will produce a set of discriminative features from measured sensor signals. For instance, static and dynamic activities should produce different signal strengths. Take the acceleration signal as an example, the signal magnitude area (SMA) calculated by the acceleration magnitude summed over three axes within each window has been found especially effective to distinguish static activities from dynamic activities, such as sitting and walking. Machado et al., 2015 and Hassan et al., 2018 use SMA and other features to improve the recognition accuracy of dynamic activities. Studies also show that Standard deviation (Std) is helpful to achieve consistently high accuracy in differentiate activities such as walking, standing, and stairs using (Laudanski et al., 2015). Some other well-applied time-domain features are median (Murao et al., 2014), variance (Mortazavi

et al., 2014), skewness (Zhang et al., 2011, Hassan et al., 2018), zero crossing rate (Suto et al., 2016), autoregressive coefficient (AR) (Hassan et al., 2018), peak-to-peak (Machado et al., 2015, Zheng et al., 2013) and so on.

Frequency-domain features are the features which are represented to describe the periodicity of signals. To produce frequency-domain features, a window of the sensor data should first be applied a transformation function, such as Fast Fourier Transform (FFT), Discrete Wavelet Transform (DWT), or Discrete Cosine Transform (DCT). The output of FFT giving is a set of basis coefficients which represent the amplitudes of the frequency components of the signal and the distribution of the signal energy. Examples of frequency-domain features based on FFT include spectral energy (Hassan et al., 2018), entropy (Hassan et al., 2018), dominant frequency (Y Wang et al., 2018, Suto et al., 2016). These FFT-derived features are reported to be beneficial to improve the recognition performance in the above-mentioned applications. Ayachi et al., 2016 demonstrate the high efficiency of DWT in their detecting and segmenting tasks for older people's daily living activities based on multiple body-worn inertial sensors. Alickovic et al., 2018 propose another automated seizure detection and prediction model based on EEG measurements. They employ wavelet packet decomposition (WPD), DWT and empirical mode decomposition (EMD) as feature extractors, and the WPD outperform the other two methods. He et al., 2009 develop a human activity system based on DCT-extracted features from acceleration data; their experimental results achieve the accuracy of 97.51%. Desai et al., 2015 also apply DCT for feature extraction on their proposed automated cardiac arrhythmia detection framework.

Most time-domain and frequency-domain features are generated from an individual channel (axis) of a sensor; such as mean and dominant frequency. On the contrary, the **hybrid features** are usually extracted from multiple sensory channels of a sensor or multiple sensors. By doing this, hybrid features implement sensor fusion through feature extraction. E.g. for the inertial sensors, several studies explore using hybrid features for HAR, e.g., the attitude angles of the wearable device, such as tilt, rotation, yaw etc.. These features are calculated by combining the values from multiple channels of an inertial sensor or multiple inertial sensors instead of a single inertial sensor, such as an accelerometer, a gyroscope or a magnetometer. Karantonis et al., 2006 and Suto et al., 2016 use the feature of tilt angle to determine the postural orientation of the user in their studies. Kundu et al., 2017 consider other hybrid features, such as pitch and roll.

The extraction of hand-crafted features depends on domain knowledge. However, hand-crafted features are easy to understand and implement. We conclude the key hand-crafted features successfully exploited in different HAR applications in Table 4, which can give strong clues for HAR tasks.

Shapelets are an important new approach for solving time series classification problems. A shapelet feature is a small subsequence extracted from the time series, which can be maximally representative of a class. Shapelet-based classification uses the similarities between a shapelet and a series as features for a classier. Shapelets are used in many tasks such as interpretable features extracting (Xing et al., 2011), gesture recognition (Hartmann et al., 2010), and gait recognition (Sivakumar et al., 2012). Since any subsequence in a time series can be a shapelet candidate, one of the challenges in this field is how to efficiently discover the shapelets and evaluate their prediction quality. Liu et al., 2015 explore the shapelet-based approaches for recognizing complex human daily activity and sport activity. They use the shapelets candidates to represent atomic activities, such as Sit, Stand and Jump, then the sequential and concurrent activities are learned from the shapelets candidates, like Relax, Cleanup, Coffee or Jump-shot. Cetin et al., 2015 present a novel technique to speed up shapelets discovery without decreasing accuracy; they use a skipping technique for pruning the additional candidates and a voting-based method to improve accuracy. Zakaria et al., 2016 present their clustering-based method on learning the shapelets from unlabelled time series. The method is tested on the diverse domains and demonstrated as highly competitive in terms of the accuracy and the discovery speed compared with the existing methods. Grabocka et al., 2016 utilize a distance-based online pruning technique to avoid measuring the significance of those similar shapelets candidates. Additionally, a supervised shapelet filtering method is employed to select the shapelets that can boost classification accuracy.

Even the speedup methods, such as clustering, pruning, and dimensionality reduction, are employed, the shapelet discovery remains computationally expensively. Hou et al., 2016 present a sparse and blocky solution by combining fused lasso regularizer and the generalized eigenvector method to transform the shapelet discovery task as a numerical optimization problem. The experimental results demonstrate their proposed method is orders of magnitudes faster than the state-of-the-art shapelet-based methods, with the comparable accuracies. However, their method is still time-consuming when dealing with the large datasets or long-time series. Also, the proposed shapelet-based methods are only compared with the other existing shapelet-based methods, and no works are seen comparing their methods with other time series classification, feature extraction or feature selection methods.

Table 4 Typical hand-crafted features used in HAR

Item	Feature title	Description	Formula (if possible)
Time-domain features	Mean (Margarito et al., 2016)	The average value of the signal over the window	$\mu = \frac{1}{T}\sum_{i=1}^{T} s_i$
	Root Mean Square (Rms) (Sani et al., 2017)	The quadratic mean value of the signal over the window	$\sqrt{\frac{1}{T}\sum_{i=1}^{T} s_i{}^2}$
	Peak-to-peak amplitude (Ptp) (Machado et al., 2015)	The difference between the maximum and the minimum value over a window	$\max\{s_1, s_2, \dots s_T\} - \min\{s_1, s_2, \dots s_T\}$
	Zero crossing rate (Czr) (Machado et al., 2015)	Rates of time signal crossing the zero value, normalized by the window length	
	Mean crossing rate (Cmr) (BanosGalvez et al., 2014)	Rates of time signal crossing the mean value, normalized by the window length	
	Signal magnitude area (SMA) (Hassan et al., 2018)	The acceleration magnitude summed over three axes within each window normalized by the window length	$\frac{1}{T}(\sum_{i=1}^{T}\|a_x(t)\| + \sum_{i=1}^{T}\|a_y(t)\| + \sum_{i=1}^{T}\|a_z(t)\|)$
	Average of peak frequency (Apf) (Janidarmian et al., 2017)	The average number of signal peak appearances in each window	
	Log-energy (Sani et al., 2017)	Log of energy	$\sum_{i=1}^{T} \log(s_i{}^2)$
	Movement Intensity (MI) (Chernbumroong et al., 2014)	Mean of the total acceleration vector over the window	$\frac{1}{T}\sum_{i=1}^{T}\sqrt{a_{xi}{}^2 + a_{yi}{}^2 + a_{zi}{}^2}$
	Variance of MI (VI) (Zhang et al., 2011)	The variance of Movement Intensity over the window	$AI = \frac{1}{T}(\sum_{i=1}^{T} MI(i) - AI)^2)$
	Averaged derivatives (Ader) (Zhang et al., 2011)	The mean value of the first order derivatives of the signal over the window	$\frac{1}{T}\sum_{i=2}^{T}\frac{s_i - s_{i-1}}{2}$
	Crest factor (Cftor) (Y Wang et al., 2016)	The ratio of peak values to the effective value over the window	$\frac{0.5(S_{max} - S_{min})}{RMS}$
	Percentiles (King et al., 2017)	$10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}$	

	Interquartile range (Interq) (King et al., 2017)	Difference between the 75th and 25th percentile	
	Autocorrelation (Autoc) (Machado et al., 2015)	The correlation between values of the process at different times	$\dfrac{\sum_{i=1}^{T-1}(s_i - \mu)(s_{i+1} - \mu)}{\sum_{i=1}^{T}(s_i - \mu)^2}$
	Pairwise correlation (Corrcoef) (Janidarmian et al., 2017)	The ratio of the covariance and the product of the standard deviations between each pair of axes	$corr_{XY} = \dfrac{\sum_{i=1}^{T}(X_i - \overline{X})(Y_i - \overline{Y}))}{\sqrt{\sum_{i=1}^{T}(X_i - \overline{X})}\sqrt{\sum_{i=1}^{T}(Y_i - \overline{Y})}}$
	Standard deviation (Std) (Laudanski et al., 2015)	Measure of the spreads of the signal over the window	$\sigma = \sqrt{\dfrac{1}{T}\sum_{i=1}^{T}(s_i - \mu)^2}$
	Coefficient of variation (Cv) (Janidarmian et al., 2017)	The ratio of the standard deviation to the mean	$\dfrac{\sigma}{\mu}$
Time-domain features	Kurtosis (Sztyler et al., 2017)	The degree of peakedness of the signal probability distribution	$\dfrac{\frac{1}{T}\sum_{i=1}^{T}(s_i-\mu)^4}{(\frac{1}{T}\sum_{i=1}^{T}(s_i-\mu)^2)^3} - 3$
	Skewness (Zhang et al., 2011)	The degree of asymmetry of the sensor signal probability distribution	$\dfrac{\frac{1}{T}\sum_{i=1}^{T}(s_i - \mu)^3}{(\frac{1}{T}\sum_{i=1}^{T}(s_i - \mu)^2)^{\frac{3}{2}}}$
	Max (Hassan et al., 2018)	The largest value in a set of data	$\max\{s_1, s_2, \dots s_T\}$
	Min (Chernbumroong et al., 2013)	The smallest value in a set of data	$\min\{s_1, s_2, \dots s_T\}$
	Median (Murao et al., 2014)	The middle number in a group of ordering numbers	$\mathrm{median}\,(s_i)$
	Mode (Chernbumroong et al., 2014)	The number that appears the most often within a set of numbers	$\mathrm{mode}\,(s_i)$
	Variance (Mortazavi et al., 2014)	The average of the squared differences from the Mean	$\dfrac{1}{T}\sum_{i=1}^{T}(s_i - \mu)^2$
	Autoregressive coefficient(AR) (Hassan et al., 2018)	Coefficients of an IIR filter, α_i	$X(n) = \sum_{i=1}^{P}\alpha_i s(n - p) + e(n)$
	Median absolute deviation(MAD) (Suto et al., 2016)	The median of the absolute deviations from the data's median	$Median_i\,(\lvert s_i - median_j(s_j)\rvert)$

Frequency-domain features	Dominant frequency (Domifq) (Suto et al., 2016)	The frequency corresponding to the maximum of the squared discrete FFT component magnitude of the signal from each sensor axis	
	Spectral energy (SpecEgy) (Hassan et al., 2018)	The sum of the squared discrete FFT component magnitude of the signal from each sensor axis, normalized by the window length	$\dfrac{\sum_{i=1}^{\lvert\omega\rvert}\lvert x_i\rvert^2}{\lvert\omega\rvert}$
	Spectral entropy (SpecEnt) (Hassan et al., 2018)	Measure of the distribution of frequency components, normalized by the window size	$\sum_{i=1}^{T/2}[P(i)\cdot\lg(P(i))]$
	The spectral centroid frequency (SCF) (Sani et al., 2017)	The estimate of the "centre of mass "of the spectrum	
Other hybrid features	Eigenvalues of dominant directions (EVA) (Zhang et al., 2011)	The relative motion magnitude along the vertical direction and the heading direction respectively	
	Averaged velocity along heading direction (AVH) Zhang et al., 2011)	Firstly, calculating the averaged velocities along y and z axes over the window, and then Computing the Euclidean norm of those two velocities	
	Pitch, yaw, roll features (Gjoreski et al., 2011b, Kundu et al., 2017)	The features extracted from the attitude values of an Inertial Measurement Unit	

2.5.2 Automatically learned features (deep features)

The second feature representation technique in current HAR applications is using deep learning techniques. Deep learning can automatically learn features from raw sensor data with less human effort, which optimizes parameters layer-by-layer following the principle that the decoded output should be equal to the input (Wang et al., 2017). The automatically learned features from deep networks are also called deep features or deep extracted features. Deep features are developed and applied in recognition tasks to improve performance (Hammerla et al., 2016, Hannink et al., 2017). For example, Ronao et al., 2016 use a deep convolutional neural network (CNN) for human activity recognition. The network they propose automatically extracts useful features from the raw data. They also investigate the effect of the performance of the extracted features from different layers on the increasing number of feature maps. The authors state their proposed network provides a way to automatically extract robust features without the requirements of pre-processing and time-consuming on feature hand-crafting. Zeng et al., 2014 propose a CNN-based feature extraction appraoch. Their experimental results indicate the extracted local dependency and scale invariant characteristics from the acceleration time series outperforms the state-of-the-art approaches.

Panwar et al., 2017 design a CNN-based framework for the recognition of three fundamental movements of the human forearm performed in daily life. Their framework learns features from the wrist-worn acceleration data. Their experimental results present the better performance of the proposed framework compared with other existing conventional methods. However, the authors do not give the details about what specific hand-crafted features they use for the conventional methods. Sani et al., 2017 also report that the automatically learned features outperform the hand-crafted features in their work. They compare the former with the latter from the time domain, frequency domain, FFT and Discrete Cosine Transform (DCT) separately. DCT performs best on the Thigh data and deep features outperform DCT slightly on the Wrist data. Whilst, their experimental results do not answer a key question whether the deep features they used can beat the combination of all the hand-crafted feature sets they use instead of beating the feature subset separately.

Some other studies explore combining hand-crafted features and deep features for HAR. Plötz et al., 2011 propose an RBM-based feature learning approach to discover universal features for activity recognition. Their experimental results based on four publicly available AR datasets indicate that combining the deep learning features with the hand-crafted features outperform other classical approaches. The results in Kashif et al., 2016 show that adding hand-crafted features to the raw data can help improve the detection accuracy of deep

convolutional neural networks for tumour cells in histology images. Meanwhile, there are some other studies giving certain interesting findings in similar fields, e.g., the experimental results in Khan et al., 2016 indicate that the hand-crafted features outperform the deep learned features in medical images. Song et al., 2016 use both video and wearable sensor data to tackle the egocentric activity recognition problem. They propose multi-stream CNN and Long short-term memory (LSTM) deep architectures to learn features from video and sensor data respectively. Experimental results indicate their proposed methods do not perform better than the hand-crafted features used in their work. They explain that this is due to that the amount of training data for their deep networks is small. Collectively, feature representation or extraction is a crucial step in HAR process. The problem of feature learning could depend on a task at hand. We produce Table 5 which summaries the advantages and disadvantages of hand-crafted features and automatically learned features based on the abovementioned studies.

Table 5 Comparison of hand-crafted features and automatically learned features

Feature type	Advantages	Disadvantages
Hand-crafted Features	Easy to understand the physical meanings of the features; Extraction is efficient and easy to deploy; Work well for many HAR problems.	Domain knowledge needed; Sensor-type specific; Need further feature selection.
Automatically learned features	No domain knowledge needed; Automatically learning features from raw data; Features are more robust and generalized.	Lots of computing resources; Parameters are difficult to adjust; The learned features are less interpretable.

2.6 Feature dimensionality reduction and feature selection

More features carry richer information, which is beneficial for improving classification performance. Feature dimension, especially for the hand-crafted features, extracted from the time, frequency or hybrid domains, becomes very high in most HAR tasks. The initial set of features can be redundant or too large to be manipulated; this could cause higher computation cost, low learning efficiency and overfitting on unseen data. Appropriate feature dimensionality reduction and feature selection can be applied in this regard to facilitate more accurate and faster learning, improving generalization and interpretability.

2.6.1 Feature dimensionality reduction

Feature dimensionality reduction is one of the two ways to address the above described issues, which reconstructs features to replace the original features by producing linear or nonlinear combinations of the input

in an unsupervised way, such as Prominent Component Analysis (PCA) (He et al., 2009), Kennel PAC (kPCA) (Hassan et al., 2018), Autoencoder (Wang, 2016), Sparse filtering (Ngiam et al., 2011) and so on. PCA is one of the well-known dimensionality reduction methods. The basic idea behind PCA is to find the optimal projection that linearly transforms the original features into a new feature space in the variance sense (Yang et al., 2012). The variables, which are ranked according to their variance (from largest to lowest) in the new feature space, are called principal components. The principal components that contribute to very high variance are preserved. kPCA finds the optimal nonlinear transformation of data, which maps the input features into a higher-dimensional feature space through a kernel function (e.g., radial basis function (RBF) kernel); followed by a typical PCA (Wu et al., 2007). PCA family are good at seeking the best representative data projection. However, it may not work well since PCA does not consider any difference in classes. Unlike PCA, Linear Discriminant Analysis (LDA) projects the original features into a new space of lower dimension by maximizing the between-class separability while minimizing their within-class variability (Uray et al., 2007). The nonlinear extension of LDA is Kernel LDA (kLDA) which performs LDA in the feature pace mapped by a nonlinear kernel function (Schölkopf et al., 1998). Hassan et al., 2018 propose a smartphone inertial sensor-based system for human activity recognition. The hand-crafted features, including mean, median, coefficients, etc., are further processed by kPCA and LDA for dimension reduction. The comparison studies show the superiority of their proposed approach.

An autoencoder network can learn a lower-dimensional representation of input data by minimizing the mean squared error between the input and the output (ideally, the input and the output are equal) (Van Der Maaten et al., 2009). An autoencoder consists of two parts, namely encoder, and decoder. The encoder aims to compress the original input data into a low-dimensional representation; the decoder tries to reconstruct the original input data based on the low-dimension representation generated by the encoder. As a result, the autoencoder is widely used to reduce the data dimension. These years, the autoencoder and its extensions demonstrate a promising ability to learn meaningful features from data for activity recognition (Chen et al., 2017, Gu et al., 2015, Chikhaoui et al., 2017). Sparse filtering is an unsupervised feature learning algorithm designed to learn features which are sparsely activated without having the need to model the data's distribution (Ngiam et al., 2011). For each sample in feature space, only a small subset of features is activated to achieve population sparsity; each feature is only activated on a small subset of the samples to reach lifetime sparsity, and features are roughly activated equally often to attain high dispersal. Hahn et al., 2015 present a neural network framework by

combining sparse filtering model and locally competitive algorithms to demonstrate their network's ability to classify human actions from the video. Raja et al., 2015 propose a feature extraction method based on deep sparse filtering to obtain robust features for unconstrained iris images. Other dimensionality reduction methods in HAR can be found from Álvarez-Meza et al., 2017, Peng et al., 2017, and Biagetti et al., 2017.

2.6.2 Feature selection (FS)

FS techniques, different from the dimensionality reduction techniques (such as PCA), select a subset from a feature set without altering the original representation of the features (Guyon et al., 2003). Thus, the selected features preserve the original semantics of the original features. An efficient feature selection can eliminate redundant features, simplify the model construction, provide the advantage of interpretability and enhance generation performance. A wide variety of feature selection methodologies have been proposed and applied in HAR. These methods can be classified into three groups based on their relationship with the inductive learning method for model construction, i.e., filter, wrapper and embedded.

The filter methods are those FS algorithms which filter out irrelevant features by evaluating the relevance of a feature to the output using certain criteria, such as correlation, distance, information, consistency, similarity and statistical measures (Gheid et al., 2016, Dessì et al., 2015). A filter algorithm first ranks the original features based on its criteria, then selects the features with higher rankings. Filter methods are independent of any classifiers, thereby being more efficient. The typical examples of filter methods are Relief (Gupta et al., 2014), Correlation-based Feature Selection (CFS) (Hemalatha et al., 2013), Mutual information (MI)-based feature selection methods (Cang et al., 2012), Canonical Correlation Analysis (CCA) (Kaya et al., 2014), etc. MI-based feature selection methods are a big family in filter methods; the algorithms in this family exploit the filter criteria based on MI which carries a correlation between features. MI and its extensions include mRMR (Peng et al., 2005), Joint Mutual Information (JMI) (Bennasar et al., 2015), Conditional Mutual Information Maximum (CMIM) (Gao et al., 2016), Double Input Symmetrical Relevance (DISR) (Meyer et al., 2006) and so on. Whilst, MI-based feature selection (FS) methods share a common problem, i.e., in some ways, it ignores the complementarity within a feature set or between features and the label since MI considers the correlation in pairs. Unlike MI, CCA measures the linear relationship between two multidimensional by maximizing the correlation coefficients between them. CCA can be used as a feature selector. CCA and its extended FS algorithms include LSCCA (Kursun et al., 2011), DCCA (Andrew et al., 2013), MCR-CCA (Kaya et al., 2014), etc.

The wrapper methods select a subset of features with the most discriminating properties by using certain classifiers to evaluate the quality of a candidate feature, e.g. SVM (Bolón-Canedo et al., 2013) and neural networks (NNs) (Kabir et al., 2010). Given a predefined classifier, a typical wrapper goes through the following process: 1) search a subset of features; 2) evaluate the selected feature set by the performance of the predefined classifier; 3) the process repeats until when the estimated accuracy of adding any feature is less than the estimated accuracy of the feature set already selected. The wrapper methods consider the features dependency and the interaction with a chaffier, thereby tending to offer a better result. While the wrapper methods are computationally expensive since performance assessments with a classifier are generally done using cross-validation (Wang et al., 2005). Thus, the wrapper methods are rarely used.

The embedded methods tend to take advantage of the merits of filter and wrapper methods by integrating feature selection into model learning (Li et al., 2017). This can be implemented by regularization techniques which introduce additional constraints (feature coefficients) into the optimization (minimizing fitting errors) simultaneously. The most widely used embedded methods are Lasso (Li et al., 2017) and Ridge regression (Liu et al., 2015). LASSO, i.e., ℓ_1-norm regularization, has the property for feature selection, which can force some feature coefficients to become smaller or exactly zero. And the features with large feature weights can be selected. Li et al., 2017 introduce group Lasso into their proposed distributed feature selection method to reduce the high dimensionality of data in the genetic study of Alzheimer's disease. Similarly, Ridge performs l_2-norm regularization for feature selection (Huang et al., 2015).

Other feature selection methods, such as sparse representation, can refer to the works in Subrahmanya et al., 2010, Liu et al., 2016 and Chu et al., 2013. There is no rigorous boundary between feature dimensionality reduction and feature selection; research continues to support the claim that there is not a "best method" for all tasks (Gui et al., 2017). The choice of the best feature set is usually with the aid of feature selection techniques or empirical evaluation of different combinations of features (Sani et al., 2017).

2.7 Classification algorithms

Classification process must be done to recognize human activities. The role of classification is to interpret the input features and give a prediction of the observations (the activity) (Alpaydin, 2014). In terms of classification algorithms used for HAR, current techniques can be categorized into two types: conventional classification algorithms and deep learning algorithms. The conventional classification algorithms attempt to build a complete description of the input with a probabilistic model such as a Bayesian network or model the

mapping from inputs (features) to outputs (activity labels) such as SVM (Chen et al., 2012). The features used by conventional classification algorithms can be the hand-crafted or automatically learned features. Deep learning algorithms are the representation-learning methods with multiple layers of representation starting from the raw data (LeCun et al., 2015). Thereby, the features can be learned automatically through the network simultaneously with the process of modelling. The features used by deep learning algorithms can also be hand-crafted features.

2.7.1 Conventional classification algorithms

From flowchart A shown in Fig.2, the features derived from raw sensor data are then fed to different classification algorithms for models constructing to classify data (e.g., the activities under consideration for HAR). The conventional classification algorithms in Fig.2 are generally categorized into two types: supervised and unsupervised. Supervised algorithms deal with labelled data and unsupervised algorithms draw inferences from datasets consisting of unlabelled input data. Supervised algorithms use training datasets to build models and test datasets to validate the models. Supervised classification is a very productive field; a large number of efficient and well-known algorithms come under this category. Some well-performed and well-known supervised algorithms are like Support Vector Machines (SVMs) (Mehrang et al., 2017), Artificial Neural Network (ANN) (Khan et al., 2014), Naïve Bayes (NB) (Mortazavi et al., 2014), Decision trees (DT) (Mortazavi et al., 2014), k-Nearest Neighbours (kNN) (Adaskevicius, 2014), Multiplayer Perceptron (MLP) (Bayat et al., 2014), Random forest (RF) (Pavey et al., 2017), etc. Atallah et al., 2011 present a framework investigating on the sensor placement and the corresponding relevance for activity recognition. They use kNN with different values of k to assess the effect of outlier points and a Bayesian classifier to model the data. Janidarmian et al., 2017 conduct a comprehensive comparison among 293 different classifiers, including DT, SVM, kNN, NB, etc., to find the best predictive model for diverse human activities. They first create the most complete dataset focusing on acceleration data and do an extensive feature extraction on data. PCA is then used for feature dimensionality reduction. The averaged accuracy achieves 96.44 ± 1.62% with k-fold cross-validation and 79.92% ± 9.68% with subject-independent cross-validation. Experiment results demonstrate that kNN and its ensemble methods show stale results over different situations, followed by ANN and SVM. The authors conclude that the determination of parameters values in each classifier can have a significant impact on the classifier's performance. They also state that certain factors, such as sensor position on body, clothing, body shape and accidental misplacements, hinder building a solid model for different activities. Mehrang et al., 2017

investigate activity monitoring using a single wrist-worn device that is equipped with an optical heart rate sensor and a triaxial accelerometer. The authors apply RF and SVM to classify a variety of home-specific activities (sitting, standing, household, and stationary cycling) performed by 20 male participants. Results of leave-one-subject-out cross-validation show 89.2% and 85.6% average accuracies from RF and SVM, respectively.

In unsupervised learning, all the sensor data are passed to the algorithm which automatically identifies a certain number of states or data clusters, each of which may correspond to a particular activity. The most common unsupervised learning method is cluster analysis, which is used for exploratory data analysis to find hidden patterns or grouping of data. The clusters are modelled using a measure of similarity which is defined upon metrics such as Euclidean or probabilistic distance. Typical unsupervised learning algorithms include k-Means (Kwon et al., 2014), Gaussian mixture models (GMM) (Kwon et al., 2014), Hidden Markov models (HMM) (Uslu et al., 2013). Mannini et al., 2011 propose a cHMM-based sequential classifier for physical activity recognition, which is indicated to outperform the GMM classier they use for the same data (99.1% vs. 92.2%). Kwon et al., 2014 present an unsupervised learning method using a smartphone sensor to overcome the needs of generating training dataset and a number of activities extending in previous studies. Experimental results demonstrate the hierarchical clustering attains above 90% accuracy when k is unknown. Their proposed approach provides a new way of automatically selecting an appropriate value of k without the generating training datasets by hand.

Some other studies combine different classification algorithms to cope with the limitations of them. Chernbumroong et al., 2015 explore combining MLP, RBF, and SVM classifiers and use GA to find the optimal combination between classifiers. Reiss et al., 2015 propose a confidence-based boosting algorithm. Experimental results indicate their proposed method significantly outperforms other boosting algorithms on most of the benchmark datasets they used and especially for larger and complex classification tasks.

2.7.2 Deep learning algorithms

The majority of the conventional classification algorithms rely on hand-crafted features (Flowchart A in Fig.2). Recent years have witnessed an area of machine learning techniques for HAR, e.g., deep learning-based networks, including CNN (Panwar et al., 2017), RNN (Hammerla et al., 2016), DBN (Hassan et al., 2018), RBM (Plötz et al., 2011), etc. Deep network can both learn deep features from raw sensor data and perform classification simultaneously (Jindong Wang et al., 2017), as shown in Flowchart B in Fig.2. Many studies have showed the superior performance of deep learning in HAR. Lane et al., 2015 investigate the question of whether

deep learning techniques can address the accuracy, robustness and efficiency on mobile sensing. The authors apply DNN, DT and GMM on activity, emotion and speaker recognition sensing tasks. Experiment setup considers the aspects of feasibility, scalability, cloud partitioning and so on, and their results provide some critical needs of the widespread adoption of sensing. Panwar et al., 2017 present a CCN-based generalized model for the recognition of three fundamental movements collected from a single wrist-worn accelerometer sensor. The comparison study between their presented method and SVM, K-means, LDA demonstrate the former outperforms. Also, their CNN-based method can handle both the feature engineering and classifying. But the authors do not give a clue whether they use delicate hand-crafted features on the latter classifiers or only pick some hand-crafted features at random. Um et al., 2017 propose a 7-layer CNN structure for augmentation of wearable data for Parkinson's disease monitoring. Ignatov, 2018 present a CNN-based deep network for online human activity recognition; their experimental results show the CNN augmented with statistical features produce a significantly-improved performance. They also demonstrate their proposed shallow architecture can be executed on mobile phones in real time. Ravi et al., 2016 also present an efficient implementation on mobile phones and the network they used is a shallow CNN structure. Suto et al., 2017 mention in their other study that a simple ANN can perform better than complex CNNs in HAR, since they believe CNN can conduct feature extraction itself whereas the CNN may not substitute the feature extraction stage in conventional techniques. Collectively, how to effectively combine hand-crafted features, automatically learned features, conventional classification algorithms, and deep learning algorithms are still worth investigations. Based on the discussions above, we summarise the characteristics of conventional and deep learning classification algorithms shown in Table 6.

3 Other two sensor modalities

3.1 Ambient sensor-based HAR (ASHAR)

Wearable sensor-based systems discussed in Section 2 have achieved wide applications in HAR due to the ease of deployment and use, low-cost and satisfied performance (Lara et al., 2013, Cornacchia et al., 2017). However, WSHAR can only provide the recognition of specific activities without giving the ambient context. Typical ambient sensors can instead provide rich contextual information relating to human daily activities, and

Table 6 Comparison of conventional and deep learning classification algorithms

Conventional	Deep learning

Features	Hand-crafted Dependent on domain knowledge	Automatically learned Independent on domain knowledge
Feature selection	Needed	No need Data pre-processing for deep networks are challenging
Model building	Model structure of a specific classifier is relatively fixed	No universal deep networks for the tasks at hand
Parameters setting and time cost	Parameters are easy to determine, comparatively takes much less time to train	A high number of hyper parameters are needed to tune, that training them takes longer

ambient sensor-based HAR (ASHAR) systems have also been widely used in HAR (Wilson et al., 2005, Tunca et al., 2014, Luo et al., 2017). This paper pays more attention to WSHAR. Therefore, the survey on other sensor modalities in this section is more compact compared to WSHAR. ASHAR systems identify human activities from the environment which is augmented with a variety of sensors, such as a door with a switch sensor, a kettle with object tags, a fridge with contact sensors, a floor with pressure sensors, a room mounted with motion sensors, etc., these sensors provide the user's contextual information where they perform activities (Debes et al., 2016, Mehr et al., 2016, Tunca et al., 2014). A wide range of ambient sensors are available and have been exploring for HAR, including cameras, light sensor, reed switch sensor, RFID, PIR, temperature, flow sensor, pressure sensor, etc. We summarise the most widely used ambient sensors in Table 7. These sensors have enabled of monitoring of daily life with somewhat general tasks.

3.1.1 Typical ambient sensor-based HAR

Typical ASHAR systems here refer to the ASHAR systems without using cameras as sensors, which detect users' activities by detecting if the user contacts the object attached with ambient sensors or by identifying whether the user enters the viewing range of one specific ambient sensor. For example, Tunca et al., 2014 develop an Ambient Assisted Living (AAL) system to infer the users' health and wellbeing status. A high number of sensors, including contact sensors, IR (infrared) receivers, sonar sensors, etc., are deployed in real environment settings. Kushwah et al., 2015 present a multi-ambient-senor framework for indoor activity recognition. Their work focuses on dealing with the difficulty of identifying the events that occur in the same context where the same set of sensors are activated during the occurrence. The authors use two smart home datasets in their experiments; one house is equipped with 14 digital sensors, such as toilet flush sensors, doors, refrigerator, and cupboards location sensors, with five different activities collected, including Drink, Dinner,

Breakfast and so on; the other house is equipped with 21 sensors, with 15 activities recorded including Toileting, Showering, Drink, Brush teeth and so on.

Luo et al., 2017 propose a framework to solve the problem of the simultaneous tracking and activity recognition (STAR). They deploy the ceiling-mounted PIR sensor array in a room. The captured information, including location, speed, and duration, is fed to the proposed two-layer RF (Random Forest) algorithm for activity recognition. The experimental results are encouraging with the recognition accuracy of above 92% for five daily activities, i.e., walking, lying, sitting, standing and transitional activities. Yasmin van Kasteren et al., 2017 explore a routine-based approach for the interpretation of smart home sensor data, they only exploit PIR sensors and power use sensors located in the participants' bathroom, lounge, bedroom, and kitchen. They successfully record 180 days of sensor data coupled with the corresponding interview data from five participants' instrumented homes. The findings from the longitudinal data demonstrate the potential of using the routines and the variation in routine to make a real-time monitoring, reliable alerts and the satisfaction of the persons being monitored. PIR sensors are also used for gait assessment in Kaye et al., 2012, the authors use a line of ceiling-attached passive infrared motion sensors for gait speed estimation and walking speed assessment from the pattern and time intervals of sensor firings. Castro et al., 2017 present a system based on the Internet of Things (IoT) to HAR by monitoring vital signs remotely. The system is successfully implemented with a 95.83% success ratio for four pre-established categories (lie, sit, walk and jog).

From the ASHAR studies given above, we can see that HAR systems deployed with typical ambient sensors are less obtrusive because the users do not need to wear any sensors. Whilst, these systems normally deploy a high number of ambient sensors at fixed locations in the environment; this will cause poor flexibility and complex sensor deployment. Also, ASHAR works in a limited area, which is usually less capable of identifying delicate actions (Debes et al., 2016, Mehr et al., 2016, Tunca et al., 2014).

3.1.2 Camera-based HAR (CHAR)

The CHAR is an active field in computer vision. There are a variety of studies on activity recognition by cameras, in which visual information acquired from the cameras mounted in fixed locations inside a building is utilized to match with the features extracted from action labels for activity recognition (Jalal et al., 2014, Jalal et al., 2017). This paper sees CHAR as ASHAR, since most CHAR systems deploy the cameras in the environment. For example, Bian et al., 2015 propose a robust fall detection approach by analysing the key joints tracked from a single depth camera. Khan et al., 2011 use one single camera to recognize six different

abnormal activities (headache, chest pain, forward fall, faint, backward fall and vomit). Binary silhouettes instead of depth silhouettes are extracted to minimize the privacy at the price of failing to distinguish different body parts. Jalal et al., 2017 present a depth video-based novel method using robust multi-features and embedded Hidden Markov Models (HMMs), with the aim of providing a health care monitoring system to support independently living for older people. The multi-features are extracted from human depth silhouettes and joint body parts information. Experimental results demonstrate the significant recognition performance and potential applications for older and sick people.

Due to the advances in 3D depth cameras, Kinect sensors (typically including an infrared camera, infrared projector and microphone array) are deployed to detect the person's full-body motion, facial recognition, voice recognition, and so on. Mohamed et al., 2012 develop a wireless sensor-based smart home, they explore Kinect sensors monitoring an older person or disabled person. Stone et al., 2015 propose a two-step approach to detect falls for older people living at home by utilizing the Microsoft Kinect sensors. Phillips et al., 2017 use Kinect sensors not only for gait change prediction but also the occurrence of future falls. They also process the Kinect depth images as silhouettes to protect privacy and embed the Kinect sensor on a small shelf above the front door to maximize the camera's view of activity. Kinect sensor systems hold promise for unobtrusively monitoring while maintaining privacy and eliminating the burden of additional monitoring procedures. Deploying a Kinect sensor set in each room at home for daily activity recognition is also less affordable.

Collectively, the significant advantage of camera-based monitoring systems is the contactless observation. And the rich information from images and videos is capable of detecting verified activities (Mabrouk et al., 2017). Whilst, sophisticated algorithms are normally needed to cope with arbitrary views of the pictures captured from cameras or complex contexts. This will cause huge time consumption. Meanwhile, it is difficult and less feasible to install cameras in all the places where older people are active. The recognition accuracy of CHAR systems may decrease due to variable lighting and other disturbances (Z Wang et al., 2017). Also, the privacy concerns cannot be ignored, although the researchers have been trying to minimize privacy by using the mini-dome and integrated cameras or exploring silhouettes instead of real pictures for activity recognition. CHAR systems are therefore more suitable for an emergency, a public safety surveillance, or scheduled meetings, instead of home-based daily monitoring.

3.2 Hybrid sensory-based HAR (ASHAR)

A HAR system normally uses a single sensor modality, i.e., wearable or ambient alone. Each sensor modality has its own strengths and limitations (as discussed in Section 2 and Section 3.1), and single sensor modalities sometimes cannot well cope with complex situations in practice. This lays the foundation for exploring hybrid sensory HAR systems. Different sensor modalities offer diverse information and varied performances for specific tasks. For example, cameras deliver precise and direct information while coupled with privacy issues or working in a constrained space defined by the camera position and settings; ambient sensors (such as the temperature or light sensor) can provide important contextual information, whilst this can only give limited information for activity detection; door switches and other binary sensors are inexpensive and easy to install, but the captured ambient information is simple and limited to detect high-level activities; the accelerometer, the gyroscope, and other wearable sensors are miniature-sized and can be flexibly worn on body to capture sufficient motion-related information. However, they cannot provide the contextual information and suffer the problem of arbitrary data caused by activities. Consequently, it is inappropriate to say which sensor modality is the best in an oversimplified way since different systems carry varied strengths and technologies targeting different applications unless we limit the task in a very specific range. Meanwhile, it is obvious that the combination of different sensor modalities can capture rich information about human activities. The following sections look into some studies which combine different sensor modalities for HAR.

3.2.1 CHAR/Audio plus WSHAR

Pansiot et al., 2007 present a sensor-fusion-based framework, in which an ear-worn accelerometer and a vision sensor installed in the environment are combined to improve classification accuracy. Hayashi et al., 2015 investigate the combination of environmental sound and acceleration data using DNN for HAR. Experimental results demonstrate the effectiveness of their proposed method with an accuracy rate of 91.7% for nine different daily activities. Liu et al., 2014a propose a hybrid sensor modality framework based on the probabilistic HMM classification for hand gesture recognition. Their framework fuses the data from an inertial sensor and a Kinect depth sensor. Their experimental results show that the accuracy can reach 93% after the data fusion while the performances of using the inertial sensor and the vision depth sensor individually are only 88% and 84%, respectively.

3.2.2 ASHAR plus WSHAR

StikicHuynh et al., 2008 investigate the feasibility of integrating RFID into wearable accelerometers on the wrist when detecting users' daily activities. Their experimental results present significantly improved recognition accuracy after sensor fusion. They utilize the number of activations from infrared sensors plus features extracted from the acceleration data as the input of the classifiers when combining the two-source data. Take active learning with 12.5% labelled data as examples in the study, the corresponding results are 60.6% $\pm$ 2.3%, 42.3% $\pm$ 2.1% and 64.2 $\pm$ 1.9%, respectively, for acceleration, infra-red data and the combined data. Roy et al., 2016 propose a hybrid approach to detect complex daily activities for multiple-inhabitant smart context by using wearable and ambient sensors, i.e., phone-carried inertial sensors and location measurement sensors. Experimental results on two separate smart home datasets demonstrate that their proposed method achieves the accuracy of 70%, which is improved by 30% compared to pure smartphone-based solutions. Y Wang et al., 2018 propose a hybrid sensory-based HAR system, which provides a more comprehensive and accurate activity monitoring for older people by combining the wrist-worn sensors and ambient-mounted PIR sensors.

3.2.3 CHAR plus ASHAR plus WSHAR

Diethe et al., 2017 introduce using Bayesian models to tackle the challenges of fusion of heterogeneous sensor modalities. The multiple-sensor-modality data, including environmental data from PIR sensors, accelerometer data, and video data, are collected in the HealthCare in the Residential Environment SPHERE house (Diethe et al., 2014). The authors summarize that their proposed approach can identify the modalities for each particular activity and the features relevant to the activity simultaneously. Also, the results show how the approach fuses and separates the tasks of activity recognition and location prediction. Nakamura et al., 2010 present a collective framework which can monitor a user's location and vitals (heart rate or blood pressure) by synchronizing wearable and ambient sensors.

3.2.4 Data fusion between different sensor modalities

Data fusion from different sensor modalities in hybrid sensory systems is found in different ways. For example, Liu et al., 2014a fuse the data from inertial sensors and vision depth for gesture recognition by feeding the fused data to HMM classifier after synchronization. This is data-level fusion. In the work by Pansiot et al., 2007, the data independently obtained from the ear-worn accelerometer and the wall mount camera are pre-processed as features before they are fed to a Bayesian classifier, this is feature-level fusion. Similarly, StikicVan Laerhoven et al., 2008 use the number of activations of infrared sensors plus features extracted from

the acceleration data as the input of the classifiers when combining the two-source data. In Liu et al., 2014b, data from differing modality sensors are fed to a multi-HMM classification framework for hand gesture recognition. Each classifier generates its own likelihood probability and the maximum of which is considered to be the recognized gesture. This is decision-level fusion. How to fuse the data from multi-sensor modalities also depends on the task purpose of a hybrid system, and it is worth investigating at different levels with diversified approaches. Following the above discussion, we summarise the three sensor modalities in Table 7.

4 Performance evaluation and application of HAR

4.1 Performance evaluation and criteria

Evaluation of the performance of a HAR system is also crucial. Two typical approaches are normally found applied in HAR applications through literature review, i.e., k-fold-cross validation (Shinmura, 2014) and leave-one-subject-out (Vehtari et al., 2017). The k-fold cross validation is a procedure used to estimate the performance of the model on unknown data (James et al., 2013). The procedure 1): shuffles the dataset available randomly, 2): then splits the dataset into k folds of approximately equal size; 3): for each unique fold, take the fold as a hold out as the test data set; take one fold from the k-1 folds as the validation data set and the remaining k-2 folds as the training data set; 4:) fit the model on the training set and evaluate it on the valuation set; 5:) test the model with the highest evaluation score and discard the other models; and the test conducts k times. The results of a k-fold cross-validation run are often summarized with the mean of the k times' test (Kuhn et al., 2013). In practice, the k value must be chosen, for example, k is set as 2 in Hu et al., 2014, 3 in ChavarriagaBayati et al., 2013, 5 in Hemalatha et al., 2013, 8 in Kreil et al., 2014, and 10 in Nam et al., 2013. The value for k is common to fix to 5 or 10 since these values have been shown empirically yielding a model performance estimate with low bias and a modest variance (James et al., 2013, Biswas et al., 2014, Ignatov, 2018). When k equals the number of subjects, the k-fold cross-validation is exactly the leave-one-subject-out cross-validation (Liu et al., 2012), which means the models are trained on the data for all subjects except one in one round, and the data from the left-out subject is used for testing. This process is repeated for each subject, and the averaged result across all the subjects is the final result (Biswas et al., 2014, Gupta et al., 2014).

Classification accuracy is the most commonly adopted performance criterion in HAR. Meanwhile, there exist other measures providing different views to understand a classification model especially for unbalanced data (Patil et al., 2013). And these criteria can be calculated from a confusion matrix. Confusion matrix, also

Table 7 Summary of sensor modalities in HAR systems

Sensor modality	Description	Sensor examples	Case study	Advantages	Disadvantages
WSHAR	Recognizing human activities by mining the informative data from wearable sensors	Accelerometer, gyroscope, heart rate, etc., built in a smartphone, band, watch, garment or other devices	Laudanski et al., 2015 Sztyler et al., 2017	Miniature-sized, low-cost, flexibly worn on body, capture motion-related information	Cannot provide the contextual information, suffer the problem of arbitrary data caused by activities
ASHAR	Inferring human activities from the sensors that are normally fixed in the environment	Surveillance camera	Phillips et al., 2017 Jalal et al., 2017	Camera can give precise and direct information	Privacy issues, expensive, working in a constrained space
		PIR, RFID, contact sensor, temperature sensor, humidity sensor etc.	Luo et al., 2017 Tunca et al., 2014 Mehr et al., 2016	provide important contextual information, less obtrusive	Limited information and working space, complex sensor deployment
HSHAR	Combining WSHAR and ASHAR for HAR	Combination of vision and accelerometers, fusion of PIR sensors and accelerometers, etc.	Hayashi et al., 2015 Diethe et al., 2017 Nakamura et al., 2010	Capture rich information and use the strengths of different sensor modalities	Complex system structure and high cost, data fusion and synchronization

known as an error matrix, is a specific matrix that allows visualization of the performance of a classification (James et al., 2013). Each row in a confusion matrix represents the instances in an actual class while each column of the matrix represents the instances in a predicted class. The element M_{ij} in an $M_{n \times n}$ matrix is the number of instances from class i that is recognized as class j actually. M_{ii} represents the number of instances from class i that is actually classified as class i. Therefore, some particular values or performance indexes can be calculated easily from the confusion matrix including TP (true positives), TN (true negatives), FP (false positives), FN (false negatives), accuracy, precision, F-score and so on (Nweke et al., 2018). Table 8 shows a basic two-class confusion matrix.

The accuracy is widely used as a statistical measure of how well a classification test correctly identifies a condition (Kwon et al., 2014). It is the proportion of true results (both true positives and true negatives)

Table 8 Confusion matrix

Actual class	Classified as	
	c1	c2
c1	TP	FN
c2	FP	TN

among the total number of cases examined, which is defined as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

The precision, on the other hand, is defined as the proportion of the true positives against all the positive results (both true positives and false positives), which is also used as the metrics in many applications (Murao et al., 2014).

$$Precision = \frac{TP}{TP + FP}$$

The recall, also called true positive rate, is the ratio of correctly classified positive instances to the total number of positive instances. In simple terms, high precision means that a classifier returns substantially more relevant results than irrelevant, while high recall means that a classifier returns most of the relevant results (Murao et al., 2014).

$$Recall = \frac{TP}{TP+FN}$$

F-measure, also called F-score, is a more comprehensive measure (Gjoreski and Gams 2011) compared to the aforementioned three ones, which combines the precision with the recall to compute the score and can be

interpreted as a weighted average of the precision and recall, where an F score reaches its best value at 1 and the worst score at 0.

$$F - measure = \frac{2 * Precision * Recall}{Precision + Recall}$$

Other performance indexes, including receiver Operating Characteristic Curve, i.e., ROC curve, and Area Under Curve, i.e., AUC, can also be seen in associated studies. A ROC represents a relation between Recall and false positive rate (specificity). AUC refers to the area under the ROC curve. Both ROC and AUC are insensitive to imbalanced classes. The studies use AUC and/or ROC for their performance assessment can refer to ChavarriagaSagha et al., 2013, Cheng et al., 2010, and Catal et al., 2015.

4.2 Applications of HAR in Health care

The recognition of human activity is not always the final goal. It is usually adopted as a paramount step for a wide range of applications, such as fitness systems, e-health care, interactive games, sports performance surveillance, social physical interaction, factory workers monitoring (Kon et al., 2017). The applications of HAR in assisted living mainly involve medical purposes and security concerns; the former focuses on monitoring patients with dementia, diabetes, obesity, arthritis or rehabilitation as an assistance diagnosis or treatment, and the latter highlights dealing with sports, entertainment, ADL, abnormal activities or safety.

Some typical WSHAR applications are as follows: Rodriguez-Martin et al., 2013 utilize a waist-attached accelerometer to identify the posture and posture transitions on healthy and Parkinson's Disease (PD) volunteers. Shibuya et al., 2015 use a gait analysis sensor set (including an accelerometer and two gyroscopes) for real-time fall detection. The sensor set is separately placed on the participant's upper end of the pelvis and the T4 area on the back. Hammerla et al., 2015 propose an assessment system, which can predict the disease state in PD patients by deploying a tri-axial accelerometer on each wrist of the participants. U M Khan et al., 2017 use passive Wi-Fi sensing for respiration-related activity monitoring by detecting breath rate, with the potential application of stress levels and psychological states assessment. Pourbabaee et al., 2017 focus on monitoring the patients with paroxysmal atrial fibrillation based on ECG time-series data from patient screening. Sathyanarayana et al., 2016 investigate the prediction of sleep quality by using deep learning methods based on a wrist-worn actigraphy, with the aim of exploring and improving eHealth solutions.

We summarise other popular applications in ASHAR, WSHAR and HSHAR systems in Table 9 in terms of sensors, features, classification algorithms, performance, etc.

Table 9 Review of existing works in HAR based on sensor modality

Sensor modality	Sensor placement		Sensor type	Sampling rate (Hz)	Window size	Feature	Activities (#)	# Subject (age)	Classifier	Performance	Target & Ref.
ASHAR	Ceiling		PIR arrays	15	1s	Hand-crafted	Walking, lying, sitting, standing, transitional (5)	3 (23 to 37)	RF	Accuracy:92%	Location & ADL [1]
	In room		Camera	NA[a]	NA[a]	Hand-crafted	Faint, backward fall, chest pain, headache, etc. (6)	6	HMM	Accuracy:95.8%	Abnormal activities [2]
WSHAR	One to One	Waist	Acc.[1]	40	3.2s	Hand-crafted	Walking, bending, lying, etc. (11)	31 healthy people, 8 patients	SVM-based	Sensitivity: 97% (healthy) Sensitivity: 98% (patients)	ADL & PD patients [3]
		Wrist	Acc.	50	1.28s	Deep[14] features	Lift cup to mouth, perform pouring, etc. (3)	4 (20 to 40)	CNN, K-means, LDA, SVM	Accuracy:99.8% (CNN)	Arm movements [4]
		Lower back	Acc.	20	6.4s /12.8 s	Hand-crafted	Walking, running, and cycling, etc. (20)	20 (29 ±6)	DT	Accuracy:93%	Indoor & outdoor activities [5]
	Multi to One	Wrist	Acc., Gyro.[2], Tem.[3], GPS, Humi.[4], Pressure	100 (Acc., Gyro) 5 (Pressure) 1 (others)	2s	Hand-crafted	Indoor to outdoor, lying on bed, Walking just, etc. (22)	2	DNN	Accuracy:90%	ADL [6]
		Wrist	Acc., Gyro.	50	2.56s	Deep & hand-crafted	Standing, sitting, laying down, walking, etc. (6)	30 (19 to 48)	CNN, NB, J48, SVM, ANN	Accuracy:95.75%	ADL [7]
	One to Multi	Lower limbs, ankle	EMG	1024	1.5s	Hand-crafted	Trip falls, stand-to-squat, stand-to-sit, walking, etc. (8)	3 (24 to 26)	FDA[5], FMMNN[6], GK-FDA[7], FCM[8], GK-SVM[9]	Accuracy:97.35% (GK-SVM8) Sensitivity:98.70% (GK-FDA)	ADL and falls [8]
		Wrist, thigh	Acc.	100	NA[b]	Deep & hand-crafted	Walking, jogging, sitting, etc. (6)	34 (18 to 54)	SVM, CNN, CNN-SVM, CNN-kNN	F1 score:0.85 (CNN-SVM, wrist) F1 score:0.967 (SVM, thigh)	ADL [9]
	Multi to Multi	Chest, thigh, ankle	Acc., Gyro., Mag.[10]	6	1s	Hand-crafted	Lying down, sitting, etc. (8)	11	RF, SVM, J48[12]	Accuracy:96.6%	ADL [10]
		Wrists, chest	Acc., Gyro., Tem., light, Baro.[11], HR[13], altimeter,	33 (Acc., Gyro) 1 (others)	3.88 s	Hand-crafted	Brushing teeth, feeding, wiping etc. (13)	12 (73±4.41)	SVM, MLP, RBF	Accuracy:97%	ADL [11]
HSHAR	Wrist, rooms		PIR, Acc., Gyro., Mag.	20	12.8s	Hand-crafted	Wash, Mop, Lie, Stand, Falls, Watch, Walk (17)	21 (60-75)	SVM, RF	Accuracy:98.96% (RF)	ADL [12]
	Room, pant pockets		PIR, Acc., Gyro.	80	5s	Hand-crafted	6: micro-activities 6: macro-activities	10	HMM	Accuracy: ~70 %	Smart environments [13]

Ref: Reference [a] NA: Not Applicable [b] NA: Not available [1] Acc.: Accelerometer [2] Gyro.: Gyroscope [3] Tem.: Temperature [4] Humi.: Humidity [5] FDA: Fisher Linear Discriminant Analysis [6] FMMNN: Fuzzy Min-Max Neural Network [7] GK-FDA: Gaussian Kernel Fisher Linear Discriminant Analysis [8] FCM: Fuzzy C-means algorithms [9] GK-SVM: Gaussian Kernel Support Vector Machine [10] Mag.: magnetometer [11] Baro.: Barometer [12] J48: the implementation of decision tree algorithm in WEKA (:a suite of machine learning software written at the University of Waikato) [13] HR: Heart rate [14] Deep features: automatically learned features by deep networks [1] Luo et al., 2017 [2] Khan et al., 2011 [3] Rodriguez-Martin et al., 2013 [4] Panwar et al., 2017 [5] Bonomi et al., 2009 [6] Vepakomma et al., 2015 [7] Ronao et al., 2016 [8] Xi et al., 2017 [9] Sani et al., 2017 [10] Gjoreski et al., 2011b [11] Chernbumroong et al., 2014 [12]Y Wang et al., 2018 [13] Roy et al., 2016

5 Open research problems and conclusion

5.1 Research problems

Research on HAR using different sensor modalities has made significant progress in continuous monitoring, performance improvement, computation cost reduction, practicability enhancement and many other domains (Chernbumroong et al., 2014, Jalal et al., 2017, Diethe et al., 2017). Due to the progress achieved in HAR- based assistive technologies, people's quality of life is being enhanced, especially those who may be physically or cognitively challenged. Nevertheless, concerns about HAR systems, including accuracy, robustness, user compliance, cost, intrusiveness, privacy and so on, make HAR still share many challenges.

- Determination of the sensor modality for a specific task

Ambient sensor-based systems are less obtrusive, whereas usually at the price of poor sensor flexibility and high cost (Tunca et al., 2014). The main concerns of using cameras at home for HAR are a high computation burden and privacy invasion (Jalal et al., 2017). As a promising way to realize HAR, wearable sensor-based HAR is low-cost, more flexible, and more practical for daily use (Cornacchia et al., 2017). Whereas, the complex sensor deployment on body for achieving higher performance may impede the user performing normal activities, increase the cost and cause the feeling of being uncomfortable, bulky and obtrusive. Some existing studies explore combining two or three of sensor modalities for HAR with the aim of using each modality's advantages and avoid their limitations (Roy et al., 2016). We should consider the fact that a proportion of older people who have limited knowledge of information technology can have less comfort with complex assistive technologies. Designing, implementing and optimizing a HAR system to satisfy the needs of older people who seek to live as independently as they can in the comfort of their home is a research problem to tackle.

- Challenges of using wrist-worn sensors for HAR

It is less feasible to wear sensors on multiple body parts for daily use in WSHAR outside of a laboratory setting. On the contrary, a wrist-worn watch-like device with embedded sensors is more convenient and less obtrusive for daily wearing. Also, the wrist is a promising position to produce high accuracy as most activities are associated with wrist movements (Mannini et al., 2013, Chernbumroong et al., 2014, Biswas et al., 2015, Mortazavi et al., 2015). Whilst, one of the most significant challenges for wrist worn sensors is the sensor signals (especially acceleration) suffer high within-class variance due to the similar attributes regarding wrist movements (Chernbumroong et al., 2013, Mortazavi et al., 2015), which lowers the recognition accuracy caused by some

easily misclassified activities, such as brushing teeth and eating (feeding), wiping and ironing (Chernbumroong et al., 2013). This imposes a challenge to activity monitoring using wrist-worn sensors. One way to overcome this challenge can be adding additional sensors to provide more sufficient information, the second can rely on feature learning from limited sensors, and another option can consider merging other sensor modality to relieve the requirements for wrist-worn sensors.

- Less fully using sensors (feature extraction)

It is common in WSHAR to use from one to seven and even more types of sensors for a specific task. Researchers prefer to acquire more diverse information through adding sensor types or sensor placing positions on body to improve performance (Gjoreski et al., 2011b, Cleland et al., 2013, Sztyler et al., 2017). These sensors are less fully used in some cases. For instance, a large number of studies exploit inertial sensors, i.e. accelerometer, gyroscope and magnetometer, but most of them only extract features from an individual sensor or multiple channels of a sensor, e.g., the mean of the acceleration readings along the *x*-axis, or the correlation between the x-axis and y-axis of the acceleration readings (Chernbumroong et al., 2014, Gjoreski et al., 2011b, Mortazavi et al., 2014). The studies above all employ limited feature sets from the sensors they choose. Only a handful of studies try few roll, yaw or pitch-related features (Gjoreski et al., 2011a, Montalto et al., 2015) derived from multiple inertial sensors as features for activity recognition, as shown in Table 4.

- Data fusion from multiple sensor modalities

Data fusion of information from multiple (usually two) sensor modalities can be done in three different ways: a) data -level, b) feature-level and c) decision-level, as discussed in Section 3. Data-level fusion occurs at the data level where incoming raw data from different sensor modalities s are combined (Liu et al., 2014a). Feature-level fusion involves carrying out data fusion after features are extracted from individual sensor modalities (Pansiot et al., 2007). Decision-level fusion involves fusing the decisions made by individual classifiers from the corresponding sensor modalities (Liu et al., 2014b). More effective and practical fusion mechanisms between ambient and wearable sensor modalities still need to be investigated.

- Hand-crafted features, automatically learned features, or both

Hand-crafted features have been successfully applying in HAR applications (Li et al., 2009, A Wang et al., 2016, Hassan et al., 2018). These years, deep learning approaches have been showing their superiority in automatically feature learning for HAR (Hammerla et al., 2015, Sani et al., 2017). The key advantages and disadvantages of hand-crafted features and automatically learned features are briefly summarized in Table 5.

Studies by Panwar et al., 2017 and Sani et al., 2017 report automatically learned features which perform better than hand-crafted features in their tasks. Plötz et al., 2011 and Kashif et al., 2016 present that combining hand-crafted features to the automatically learned features from raw data can help improve the detection accuracy of deep neural networks. Meanwhile, Khan et al., 2016 and Song et al., 2016 indicate that the hand-crafted features outperform the automatically learned features in their studies. Therefore, how to effectively use features for a HAR task is still challenging. To the best of our knowledge, very few researchers have investigated the performance of using deep networks learning deep features from hand-crafted features.

5.2 Conclusion

Sensor-based HAR systems have been achieving continuous progress. Each sensor modality has its own strengths and weaknesses. Camera-based HAR delivers direct and precise information about HAR under monitored, while companied with privacy concerns for daily use and constrained function space caused by camera settings and installation position. Ambient senor-based HAR offers ambient context, but which usually provide limited information about the human activity. Wearable sensor-based HAR is more flexible for long-term use and can provide rich motion information, however, which often suffer the problems, like arbitrary signal caused by the sensors worn on body parts. The hybrid sensory HAR which combines ambient and wearable sensor modalities can provide richer or complementary information from different sensors. Nevertheless, a combination of different sensor modalities can also involve the problems, such as increasing the complexity of the system and costs, effective data fusion between different sensor modalities. The discussion above is also summarised in Table 7.

This paper presents a survey on the wearable sensor modality centred HAR in health care, including the sensors used in HAR, the sensor placement on different body parts, the most common seen sensor platforms in HAR, activities defined in this field, data segmentation, feature learning, classification, etc. Extracting effective features for identifying activities is a critical and challenging task. For the feature learning, we survey both the commonly used hand-crafted features and automatically learned features using deep networks. Hand-crafted features are interpretable and have achieved great success in HAR. Nevertheless, there are no universal procedures for selecting appropriate features from hand-crafted features for a given human activity recognition system. Automatically learned features are obtained from raw data without any domain knowledge and can be used for classification simultaneously. Deep learning techniques have been developed and successfully applied in recognition tasks. The pros and cons of hand-crafted and automatically learned features in HAR are presented in Table 5. Meanwhile, there are some other studies giving certain interesting findings, e.g., the hand-crafted features

outperform the automatically learned features in the medical image field, or the combination of hand-crafted features with raw data produces better detection results than the results of raw intensities with a similar kind of CNN architecture. Consequently, the feature learning could depend on a task at hand. The survey also summaries the typical applications of HAR in healthcare and proposes some research challenges for further improvement.

References

Abowd D., Dey, A. K., Orr, R., & Brotherton, J. (1998). Context-awareness in wearable and ubiquitous computing. *Virtual Reality, 3*, 200-211.

Adaskevicius R. (2014). Method for recognition of the physical activity of human being using a wearable accelerometer. *Elektronika ir Elektrotechnika, 20*, 127-131.

Alickovic E., Kevric, J., & Subasi, A. (2018). Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction. *Biomedical Signal Processing and Control, 39*, 94-102.

Alpaydin E. (2014). *Introduction to machine learning*: MIT press.

Alshurafa N., Eastwood, J.-A., Pourhomayoun, M., Nyamathi, S., Bao, L., Mortazavi, B., & Sarrafzadeh, M. (2014). Anti-cheating: Detecting self-inflicted and impersonator cheaters for remote health monitoring systems with wearable sensors. In *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on* (pp. 92-97): IEEE.

Alvarez-Alvarez A., Alonso, J. M., & Trivino, G. (2013). Human activity recognition in indoor environments by means of fusing information extracted from intensity of WiFi signal and accelerations. *Information Sciences, 233*, 162-182.

Álvarez-Meza A. M., Lee, J. A., Verleysen, M., & Castellanos-Dominguez, G. (2017). Kernel-based dimensionality reduction using Renyi's α-entropy measures of similarity. *Neurocomputing, 222*, 36-46.

Andrew G., Arora, R., Bilmes, J., & Livescu, K. (2013). Deep canonical correlation analysis. In *International Conference on Machine Learning* (pp. 1247-1255).

Anguita D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2013). A public domain dataset for human activity recognition using smartphones. In *Esann*.

Anwary A. R., Yu, H., & Vassallo, M. (2017). Optimal foot location for placing wearable IMU sensors and automatic feature extraction for gait analysis. *IEEE Sensors Journal, 18*, 2555-2567.

Anwary A. R., Yu, H., & Vassallo, M. (2018). An Automatic Gait Feature Extraction Method for Identifying Gait Asymmetry Using Wearable Sensors. *Sensors, 18*, 676.

Atallah L., ElHelw, M., Pansiot, J., Stoyanov, D., Wang, L., Lo, B., & Yang, G.-Z. (2007). Behaviour profiling with ambient and wearable sensing. In *4th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2007)* (pp. 133-138): Springer.

Atallah L., Lo, B., King, R., & Yang, G.-Z. (2010). Sensor placement for activity detection using wearable accelerometers. In *Body Sensor Networks (BSN), 2010 International Conference on* (pp. 24-29): IEEE.

Atallah L., Lo, B., King, R., & Yang, G.-Z. (2011). Sensor positioning for activity recognition using wearable accelerometers. *IEEE transactions on biomedical circuits and systems, 5*, 320-329.

Attal F., Mohammed, S., Dedabrishvili, M., Chamroukhi, F., Oukhellou, L., & Amirat, Y. (2015). Physical human activity recognition using wearable sensors. *Sensors, 15*, 31314-31338.

Augimeri A., Fortino, G., Rege, M. R., Handziski, V., & Wolisz, A. (2010). A cooperative approach for handshake detection based on body sensor networks. In *Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on* (pp. 281-288): IEEE.

Ayachi F. S., Nguyen, H. P., Lavigne-Pelletier, C., Goubault, E., Boissy, P., & Duval, C. (2016). Wavelet-based algorithm for auto-detection of daily living activities of older adults captured by multiple inertial measurement units (IMUs). *Physiological measurement, 37*, 442.

Bahrepour M., Meratnia, N., Taghikhaki, Z., & Havinga, P. J. (2011). Sensor fusion-based activity recognition for Parkinson patients. In *Sensor Fusion-Foundation and Applications*: InTech.

Banos O., Damas, M., Pomares, H., Prieto, A., & Rojas, I. (2012). Daily living activity recognition based on statistical feature quality group selection. *Expert Systems with Applications, 39*, 8013-8021.

Banos O., Damas, M., Pomares, H., Rojas, F., Delgado-Marquez, B., & Valenzuela, O. (2013). Human activity recognition based on a sensor weighting hierarchical classifier. *Soft Computing, 17*, 333-343.

Baños O., Damas, M., Pomares, H., Rojas, I., Tóth, M. A., & Amft, O. (2012). A benchmark dataset to evaluate sensor displacement in activity recognition. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (pp. 1026-1035): ACM.

Banos O., Galvez, J.-M., Damas, M., Pomares, H., & Rojas, I. (2014). Window size impact in human activity recognition. *Sensors, 14*, 6474-6499.

Banos O., Garcia, R., Holgado-Terriza, J. A., Damas, M., Pomares, H., Rojas, I., Saez, A., & Villalonga, C. (2014). mHealthDroid: a novel framework for agile development of mobile health applications. In *International Workshop on Ambient Assisted Living* (pp. 91-98): Springer.

Bao L., & Intille, S. (2004). Activity recognition from user-annotated acceleration data. *Pervasive computing*, 1-17.

Barreto A., Oliveira, R., Sousa, F., Cardoso, A., & Duarte, C. (2014). Environment-aware system for Alzheimer's patients. In *Wireless Mobile Communication and Healthcare (Mobihealth), 2014 EAI 4th International Conference on* (pp. 300-303): IEEE.

Bayat A., Pomplun, M., & Tran, D. A. (2014). A study on human activity recognition using accelerometer data from smartphones. *Procedia Computer Science, 34*, 450-457.

Bennasar M., Hicks, Y., & Setchi, R. (2015). Feature selection using joint mutual information maximisation. *Expert Systems with Applications, 42*, 8520-8532.

Berchtold M., Budde, M., Schmidtke, H. R., & Beigl, M. (2010). An extensible modular recognition concept that makes activity recognition practical. In *Annual Conference on Artificial Intelligence* (pp. 400-409): Springer.

Bergmann J. H., Chandaria, V., & McGregor, A. (2012). Wearable and implantable sensors: the patient's perspective. *Sensors, 12*, 16695-16709.

Bhattacharya S., & Lane, N. D. (2016). From smart to deep: Robust activity recognition on smartwatches using deep learning. In *Pervasive Computing and Communication Workshops (PerCom Workshops), 2016 IEEE International Conference on* (pp. 1-6): IEEE.

Biagetti G., Crippa, P., Falaschetti, L., Orcioni, S., & Turchetti, C. (2017). Human Activity Recognition Using Accelerometer and Photoplethysmographic Signals. In *International Conference on Intelligent Decision Technologies* (pp. 53-62): Springer.

Bian Z.-P., Hou, J., Chau, L.-P., & Magnenat-Thalmann, N. (2015). Fall detection based on body part tracking using a depth camera. *IEEE journal of biomedical and health informatics, 19*, 430-439.

Biswas D., Cranny, A., Gupta, N., Maharatna, K., Achner, J., Klemke, J., Jöbges, M., & Ortmann, S. (2015). Recognizing upper limb movements with wrist worn inertial sensors using k-means clustering classification. *Human movement science, 40*, 59-76.

Biswas D., Cranny, A., Rahim, A. F., Gupta, N., Maharatna, K., Harris, N. R., & Ortmann, S. (2014). On the data analysis for classification of elementary upper limb movements. *Biomedical Engineering Letters, 4*, 403-413.

Bolón-Canedo V., Sánchez-Maroño, N., & Alonso-Betanzos, A. (2013). A review of feature selection methods on synthetic data. *Knowledge and information systems, 34*, 483-519.

Bonomi A. G., Goris, A. H., Yin, B., & Westerterp, K. R. (2009). Detection of type, duration, and intensity of physical activity using an accelerometer. *Medicine & Science in Sports & Exercise, 41*, 1770-1777.

Bourke A. K., Van de Ven, P. W., Chaya, A. E., OLaighin, G. M., & Nelson, J. (2008). Testing of a long-term fall detection system incorporated into a custom vest for the elderly. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE* (pp. 2844-2847): IEEE.

Bulling A., Blanke, U., & Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR), 46*, 33.

Burns A., Greene, B. R., McGrath, M. J., O'Shea, T. J., Kuris, B., Ayer, S. M., Stroiescu, F., & Cionca, V. (2010). SHIMMER™–A wireless sensor platform for noninvasive biomedical research. *IEEE Sensors Journal, 10*, 1527-1534.

Cang S., & Yu, H. (2012). Mutual information based input feature selection for classification problems. *Decision Support Systems, 54*, 691-698.

Capela N. A., Lemaire, E. D., & Baddour, N. (2015). Novel algorithm for a smartphone-based 6-minute walk test application: algorithm, application development, and evaluation. *Journal of neuroengineering and rehabilitation, 12*, 19.

Carmeli E., Imam, B., & Merrick, J. (2016). Assistive Technology and Older Adults. In *Health Care for People with Intellectual and Developmental Disabilities across the Lifespan* (pp. 1465-1471): Springer.

Castro D., Coral, W., Rodriguez, C., Cabra, J., & Colorado, J. (2017). Wearable-Based Human Activity Recognition Using and IoT Approach. *Journal of Sensor and Actuator Networks, 6*, 28.

Catal C., Tufekci, S., Pirmit, E., & Kocabag, G. (2015). On the use of ensemble of classifiers for accelerometer-based activity recognition. *Applied Soft Computing, 37*, 1018-1022.

Cetin M. S., Mueen, A., & Calhoun, V. D. (2015). Shapelet ensemble for multi-dimensional time series. In *Proceedings of the 2015 SIAM International Conference on Data Mining* (pp. 307-315): SIAM.

Chamroukhi F., Mohammed, S., Trabelsi, D., Oukhellou, L., & Amirat, Y. (2013). Joint segmentation of multivariate time series with hidden process regression for human activity recognition. *Neurocomputing, 120*, 633-644.

Chavarriaga R., Bayati, H., & Millán, J. d. R. (2013). Unsupervised adaptation for acceleration-based activity recognition: robustness to sensor displacement and rotation. *Personal and Ubiquitous Computing, 17*, 479-490.

Chavarriaga R., Sagha, H., Calatroni, A., Digumarti, S. T., Tröster, G., Millán, J. d. R., & Roggen, D. (2013). The Opportunity challenge: A benchmark database for on-body sensor-based activity recognition. *Pattern Recognition Letters, 34*, 2033-2042.

Chen G., Wang, A., Zhao, S., Liu, L., & Chang, C.-Y. (2017). Latent feature learning for activity recognition using simple sensors in smart homes. *Multimedia Tools and Applications*, 1-19.

Chen L.-L., Zhang, J., Zou, J.-Z., Zhao, C.-J., & Wang, G.-S. (2014). A framework on wavelet-based nonlinear features and extreme learning machine for epileptic seizure detection. *Biomedical Signal Processing and Control, 10*, 1-10.

Chen L., Hoey, J., Nugent, C. D., Cook, D. J., & Yu, Z. (2012). Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42*, 790-808.

Cheng J., Amft, O., & Lukowicz, P. (2010). Active capacitive sensing: Exploring a new wearable sensing modality for activity recognition. In *International Conference on Pervasive Computing* (pp. 319-336): Springer.

Chernbumroong S., Cang, S., Atkins, A., & Yu, H. (2013). Elderly activities recognition and classification for applications in assisted living. *Expert Systems with Applications, 40*, 1662-1674.

Chernbumroong S., Cang, S., & Yu, H. (2014). A practical multi-sensor activity recognition system for home-based care. *decision support systems, 66*, 61-70.

Chernbumroong S., Cang, S., & Yu, H. (2015). Genetic algorithm-based classifiers fusion for multisensor activity recognition of elderly people. *IEEE journal of biomedical and health informatics, 19*, 282-289.

Chikhaoui B., & Gouineau, F. (2017). Towards Automatic Feature Extraction for Activity Recognition from Wearable Sensors: A Deep Learning Approach. In *2017 IEEE International Conference on Data Mining Workshops (ICDMW)* (pp. 693-702): IEEE.

Chu D., Liao, L., Ng, M., & Zhang, X. (2013). Sparse kernel canonical correlation analysis. In *Proceedings of International Multiconference of Engineers and Computer Scientists*.

Cleland I., Kikhia, B., Nugent, C., Boytsov, A., Hallberg, J., Synnes, K., McClean, S., & Finlay, D. (2013). Optimal placement of accelerometers for the detection of everyday activities. *Sensors, 13*, 9183-9200.

Cook A. J., Gargiulo, G., Lehmann, T., & Hamilton, T. J. (2015). Open platform, eight-channel, portable bio-potential and activity data logger for wearable medical device development. *Electronics Letters, 51*, 1641-1643.

Cornacchia M., Ozcan, K., Zheng, Y., & Velipasalar, S. (2017). A survey on activity detection and classification using wearable sensors. *IEEE Sensors Journal, 17*, 386-403.

Dalton A., & ÓLaighin, G. (2013). Comparing supervised learning techniques on the task of physical activity recognition. *IEEE journal of biomedical and health informatics, 17*, 46-52.

Davis K., Owusu, E., Bastani, V., Marcenaro, L., Hu, J., Regazzoni, C., & Feijs, L. (2016). Activity recognition based on inertial sensors for ambient assisted living. In *Information Fusion (FUSION), 2016 19th International Conference on* (pp. 371-378): IEEE.

De P., Chatterjee, A., & Rakshit, A. (2017). Recognition of Human behavior for Assisted Living using Dictionary Learning Approach. *IEEE Sensors Journal*.

Debes C., Merentitis, A., Sukhanov, S., Niessen, M., Frangiadakis, N., & Bauer, A. (2016). Monitoring activities of daily living in smart homes: Understanding human behavior. *IEEE Signal Processing Magazine, 33*, 81-94.

Deng W.-Y., Zheng, Q.-H., & Wang, Z.-M. (2014). Cross-person activity recognition using reduced kernel extreme learning machine. *Neural Networks, 53*, 1-7.

Desai U., Martis, R. J., Nayak, C. G., Sarika, K., Nayak, S. G., Shirva, A., Nayak, V., & Mudassir, S. (2015). Discrete cosine transform features in automated classification of cardiac arrhythmia beats. In *Emerging research in computing, information, communication and applications* (pp. 153-162): Springer.

Dessì N., & Pes, B. (2015). Similarity of feature selection methods: An empirical study across data intensive classification tasks. *Expert Systems with Applications, 42*, 4632-4642.

Diethe T., Twomey, N., Kull, M., Flach, P., & Craddock, I. (2017). Probabilistic sensor fusion for ambient assisted living. *Information Fusion*.

Diethe T. R., Twomey, N., & Flach, P. (2014). SPHERE: A sensor platform for healthcare in a residential environment. In *Proceedings of Large-scale Online Learning and Decision Making Workshop*.

Farah J. D., Baddour, N., & Lemaire, E. D. (2019). Design, development, and evaluation of a local sensor-based gait phase recognition system using a logistic model decision tree for orthosis-control. *Journal of neuroengineering and rehabilitation, 16*, 22.

Filippoupolitis A., Oliff, W., Takand, B., & Loukas, G. (2017). Location-Enhanced Activity Recognition in Indoor Environments Using Off the Shelf Smart Watch Technology and BLE Beacons. *Sensors, 17*, 1230.

Fontana J. M., Higgins, J. A., Schuckers, S. C., Bellisle, F., Pan, Z., Melanson, E. L., Neuman, M. R., & Sazonov, E. (2015). Energy intake estimation from counts of chews and swallows. *Appetite, 85*, 14-21.

Gao L., Bourke, A., & Nelson, J. (2014). Evaluation of accelerometer based multi-sensor versus single-sensor activity recognition systems. *Medical engineering & physics, 36*, 779-785.

Gao S., Ver Steeg, G., & Galstyan, A. (2016). Variational information maximization for feature selection. In *Advances in Neural Information Processing Systems* (pp. 487-495).

Georgi M., Amma, C., & Schultz, T. (2015). Recognizing Hand and Finger Gestures with IMU based Motion and EMG based Muscle Activity Sensing. In *BIOSIGNALS* (pp. 99-108).

Gheid Z., & Challal, Y. (2016). Novel Efficient and Privacy-Preserving Protocols For Sensor-Based Human Activity Recognition. In *Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld), 2016 Intl IEEE Conferences* (pp. 301-308): IEEE.

Gjoreski H., & Gams, M. (2011a). Activity/Posture recognition using wearable sensors placed on different body locations. *The Fourteenth International Conference on Artificial Intelligence and Soft Computing*.

Gjoreski H., & Gams, M. (2011b). Activity/Posture recognition using wearable sensors placed on different body locations. *Proceedings of (738) Signal and Image Processing and Applications, Crete, Greece, 2224*, 716724.

Grabocka J., Wistuba, M., & Schmidt-Thieme, L. (2016). Fast classification of univariate and multivariate time series through shapelet discovery. *Knowledge and information systems, 49*, 429-454.

Gu F., Flórez-Revuelta, F., Monekosso, D., & Remagnino, P. (2015). Marginalised stacked denoising autoencoders for robust representation of real-time multi-view action recognition. *Sensors, 15*, 17209-17231.

Gui J., Sun, Z., Ji, S., Tao, D., & Tan, T. (2017). Feature selection based on structured sparsity: A comprehensive study. *IEEE transactions on neural networks and learning systems, 28*, 1490-1507.

Guo J., Zhou, X., Sun, Y., Ping, G., Zhao, G., & Li, Z. (2016). Smartphone-Based Patients' activity recognition by using a self-learning scheme for medical monitoring. *Journal of medical systems, 40*, 140.

Guo Y., He, W., & Gao, C. (2012). Human activity recognition by fusing multiple sensor nodes in the wearable sensor systems. *Journal of Mechanics in Medicine and Biology, 12*, 1250084.

Gupta P., & Dallas, T. (2014). Feature selection and activity recognition system using a single triaxial accelerometer. *IEEE Transactions on Biomedical Engineering, 61*, 1780-1786.

Guyon I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research, 3*, 1157-1182.

Hahn W. E., Lewkowitz, S., Lacombe, D. C., & Barenholtz, E. (2015). Deep learning human actions from video via sparse filtering and locally competitive algorithms. *Multimedia Tools and Applications, 74*, 10097-10110.

Hajihashemi Z., & Popescu, M. (2013). Detection of abnormal sensor patterns in eldercare. *E-Health and Bioengineering Conference (EHB), 2013*, 1-4.

Hammerla N. Y., Fisher, J., Andras, P., Rochester, L., Walker, R., & Plötz, T. (2015). PD Disease State Assessment in Naturalistic Environments Using Deep Learning. In *AAAI* (pp. 1742-1748).

Hammerla N. Y., Halloran, S., & Ploetz, T. (2016). Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv preprint arXiv:1604.08880*.

Hannink J., Kautz, T., Pasluosta, C. F., Gaßmann, K.-G., Klucken, J., & Eskofier, B. M. (2017). Sensor-based gait parameter extraction with deep convolutional neural networks. *IEEE journal of biomedical and health informatics, 21*, 85-93.

Hartmann B., & Link, N. (2010). Gesture recognition with inertial sensors and optimized DTW prototypes. In *Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on* (pp. 2102-2109): IEEE.

Hassan M. M., Uddin, M. Z., Mohamed, A., & Almogren, A. (2018). A robust human activity recognition system using smartphone sensors and deep learning. *Future Generation Computer Systems, 81*, 307-313.

Hayashi T., Nishida, M., Kitaoka, N., & Takeda, K. (2015). Daily activity recognition based on DNN using environmental sound and acceleration signals. In *Signal Processing Conference (EUSIPCO), 2015 23rd European* (pp. 2306-2310): IEEE.

He Z., & Bai, X. (2014). A wearable wireless body area network for human activity recognition. In *Ubiquitous and Future Networks (ICUFN), 2014 Sixth International Conf on* (pp. 115-119): IEEE.

He Z., & Jin, L. (2009). Activity recognition from acceleration data based on discrete consine transform and SVM. In *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on* (pp. 5041-5044): IEEE.

Hemalatha C. S., & Vaidehi, V. (2013). Frequent bit pattern mining over tri-axial accelerometer data streams for recognizing human activities and detecting fall. *Procedia Computer Science, 19*, 56-63.

Hermanis A., Cacurs, R., Nesenbergs, K., Greitans, M., Syundyukov, E., & Selavo, L. (2016). Wearable Sensor System for Human Biomechanics Monitoring. In *EWSN* (pp. 247-248).

Hexoshin. (2018). Hexoshin Smart Shirts Specifications. In (Vol. 2018).

Hou L., Kwok, J. T., & Zurada, J. M. (2016). Efficient learning of timeseries shapelets. In *Thirtieth AAAI Conference on Artificial Intelligence*.

Hu L., Chen, Y., Wang, S., & Chen, Z. (2014). b-COELM: A fast, lightweight and accurate activity recognition model for mini-wearable devices. *Pervasive and Mobile Computing, 15*, 200-214.

Huang H., Yoo, S., & Kasiviswanathan, S. P. (2015). Unsupervised feature selection on data streams. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management* (pp. 1031-1040): ACM.

Huang J.-y., & Tsai, C.-H. (2007). A wearable computing environment for the security of a large-scale factory. In *International Conference on Human-Computer Interaction* (pp. 1113-1122): Springer.

Ignatov A. (2018). Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. *Applied Soft Computing, 62*, 915-922.

Jalal A., Kamal, S., & Kim, D. (2017). A Depth Video-based Human Detection and Activity Recognition using Multi-features and Embedded Hidden Markov Models for Health Care Monitoring Systems. *International Journal of Interactive Multimedia & Artificial Intelligence, 4*.

James G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112): Springer.

Janidarmian M., Roshan Fekr, A., Radecka, K., & Zilic, Z. (2017). A Comprehensive Analysis on Wearable Acceleration Sensors in Human Activity Recognition. *Sensors, 17*, 529.

Jung S., Hong, S., Kim, J., Lee, S., Hyeon, T., Lee, M., & Kim, D.-H. (2015). Wearable fall detector using integrated sensors and energy devices. *Scientific reports, 5*, 17081.

Kabir M. M., Islam, M. M., & Murase, K. (2010). A new wrapper feature selection approach using neural network. *Neurocomputing, 73*, 3273-3283.

Kalantarian H., Alshurafa, N., Le, T., & Sarrafzadeh, M. (2015). Monitoring eating habits using a piezoelectric sensor-based necklace. *Computers in biology and medicine, 58*, 46-55.

Karantonis D. M., Narayanan, M. R., Mathie, M., Lovell, N. H., & Celler, B. G. (2006). Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE transactions on information technology in biomedicine, 10*, 156-167.

Kashif M. N., Raza, S. E. A., Sirinukunwattana, K., Arif, M., & Rajpoot, N. (2016). Handcrafted features with convolutional neural networks for detection of tumor cells in histology images. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on* (pp. 1029-1032): IEEE.

Kaya H., Eyben, F., Salah, A. A., & Schuller, B. (2014). CCA based feature selection with application to continuous depression recognition from acoustic speech features. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on* (pp. 3729-3733): IEEE.

Kaye J., Mattek, N., Dodge, H., Buracchio, T., Austin, D., Hagler, S., Pavel, M., & Hayes, T. (2012). One walk a year to 1000 within a year: Continuous in-home unobtrusive gait assessment of older adults. *Gait & posture, 35*, 197-202.

Ke S.-R., Thuc, H. L. U., Lee, Y.-J., Hwang, J.-N., Yoo, J.-H., & Choi, K.-H. (2013). A review on video-based human activity recognition. *Computers, 2*, 88-131.

Khan A. M., Tufail, A., Khattak, A. M., & Laine, T. H. (2014). Activity recognition on smartphones via sensor-fusion and kda-based svms. *International Journal of Distributed Sensor Networks, 10*, 503291.

Khan S., & Yong, S.-P. (2016). A comparison of deep learning and hand crafted features in medical image modality classification. In *Computer and Information Sciences (ICCOINS), 2016 3rd International Conference on* (pp. 633-638): IEEE.

Khan S. S., & Hoey, J. (2017). Review of fall detection techniques: A data availability perspective. *Medical engineering & physics, 39*, 12-22.

Khan U. M., Kabir, Z., Hassan, S. A., & Ahmed, S. H. (2017). A Deep Learning Framework using Passive WiFi Sensing for Respiration Monitoring. *arXiv preprint arXiv:1704.05708*.

Khan Z. A., & Sohn, W. (2011). Abnormal human activity recognition system based on R-transform and kernel discriminant technique for elderly home care. *IEEE Transactions on Consumer Electronics, 57*.

King R. C., Villeneuve, E., White, R. J., Sherratt, R. S., Holderbaum, W., & Harwin, W. S. (2017). Application of data fusion techniques and technologies for wearable health monitoring. *Medical Engineering and Physics, 42*, 1-12.

Kon B., Lam, A., & Chan, J. (2017). Evolution of Smart Homes for the Elderly. In *Proceedings of the 26th International Conference on World Wide Web Companion* (pp. 1095-1101): International World Wide Web Conferences Steering Committee.

Kreil M., Sick, B., & Lukowicz, P. (2014). Dealing with human variability in motion based, wearable activity recognition. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on* (pp. 36-40): IEEE.

Kuerbis A., Mulliken, A., Muench, F., Moore, A. A., & Gardner, D. (2017). Older adults and mobile technology: Factors that enhance and inhibit utilization in the context of behavioral health.

Kuhn M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26): Springer.

Kundu A. S., Mazumder, O., Lenka, P. K., & Bhaumik, S. (2017). Hand Gesture Recognition Based Omnidirectional Wheelchair Control Using IMU and EMG Sensors. *Journal of Intelligent & Robotic Systems*, 1-13.

Kursun O., Alpaydin, E., & Favorov, O. V. (2011). Canonical correlation analysis using within-class coupling. *Pattern Recognition Letters, 32*, 134-144.

Kushwah A., Kumar, S., & Hegde, R. M. (2015). Multi-sensor data fusion methods for indoor activity recognition using temporal evidence theory. *Pervasive and Mobile Computing, 21*, 19-29.

Kwapisz J. R., Weiss, G. M., & Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter, 12*, 74-82.

Kwon Y., Kang, K., & Bae, C. (2014). Unsupervised learning for human activity recognition using smartphone sensors. *Expert Systems with Applications, 41*, 6067-6074.

Lane N. D., & Georgiev, P. (2015). Can deep learning revolutionize mobile sensing? In *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications* (pp. 117-122): ACM.

Lara O. D., & Labrador, M. A. (2013). A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys and Tutorials, 15*, 1192-1209.

Laudanski A., Brouwer, B., & Li, Q. (2015). Activity classification in persons with stroke based on frequency features. *Medical engineering & physics, 37*, 180-186.

LeCun Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature, 521*, 436.

Lee S.-M., Yoon, S. M., & Cho, H. (2017). Human activity recognition from accelerometer data using Convolutional Neural Network. In *Big Data and Smart Computing (BigComp), 2017 IEEE International Conference on* (pp. 131-134): IEEE.

Li J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., & Liu, H. (2017). Feature selection: A data perspective. *ACM Computing Surveys (CSUR), 50*, 94.

Li Q., Stankovic, J. A., Hanson, M. A., Barth, A. T., Lach, J., & Zhou, G. (2009). Accurate, fast fall detection using gyroscopes and accelerometer-derived posture information. In *Wearable and Implantable Body Sensor Networks, 2009. BSN 2009. Sixth International Workshop on* (pp. 138-143): IEEE.

Li Q., Zhu, D., Zhang, J., Hibar, D. P., Jahanshad, N., Wang, Y., Ye, J., Thompson, P. M., & Wang, J. (2017). Large-scale Feature Selection of Risk Genetic Factors for Alzheimer's Disease via Distributed Group Lasso Regression. *arXiv preprint arXiv:1704.08383*.

Liu K., Chen, C., Jafari, R., & Kehtarnavaz, N. (2014a). Fusion of inertial and depth sensor data for robust hand gesture recognition. *IEEE Sensors Journal, 14*, 1898-1903.

Liu K., Chen, C., Jafari, R., & Kehtarnavaz, N. (2014b). Multi-HMM classification for hand gesture recognition using two differing modality sensors. In *Circuits and Systems Conference (DCAS), 2014 IEEE Dallas* (pp. 1-4): IEEE.

Liu L., Peng, Y., Liu, M., & Huang, Z. (2015). Sensor-based human activity recognition system with a multilayered model using time series shapelets. *Knowledge-Based Systems, 90*, 138-152.

Liu L., Peng, Y., Wang, S., Liu, M., & Huang, Z. (2016). Complex activity recognition using time series pattern dictionary learned from ubiquitous sensors. *Information Sciences, 340*, 41-57.

Liu M., & Zhang, D. (2016). Pairwise constraint-guided sparse learning for feature selection. *IEEE transactions on cybernetics, 46*, 298-310.

Liu P., Yu, H., & Cang, S. (2018). Optimized adaptive tracking control for an underactuated vibro-driven capsule system. *Nonlinear Dynamics, 94*, 1803-1817.

Liu S., Gao, R. X., John, D., Staudenmayer, J. W., & Freedson, P. S. (2012). Multisensor data fusion for physical activity assessment. *IEEE Transactions on Biomedical Engineering, 59*, 687-696.

Liu Y., Nie, L., Han, L., Zhang, L., & Rosenblum, D. S. (2015). Action2Activity: Recognizing Complex Activities from Sensor Data. In *IJCAI* (Vol. 2015, pp. 1617-1623).

Liu Y., Nie, L., Liu, L., & Rosenblum, D. S. (2016). From action to activity: sensor-based activity recognition. *Neurocomputing, 181*, 108-115.

Logan B., Healey, J., Philipose, M., Tapia, E. M., & Intille, S. (2007). A long-term evaluation of sensing modalities for activity recognition. *International conference on Ubiquitous computing*, 483-500.

Lorussi F., Carbonaro, N., De Rossi, D., Paradiso, R., Veltink, P., & Tognetti, A. (2016). Wearable textile platform for assessing stroke patient treatment in daily life conditions. *Frontiers in bioengineering and biotechnology, 4*, 28.

Luo X., Guan, Q., Tan, H., Gao, L., Wang, Z., & Luo, X. (2017). Simultaneous Indoor Tracking and Activity Recognition Using Pyroelectric Infrared Sensors. *Sensors, 17*, 1738.

Mabrouk A. B., & Zagrouba, E. (2017). Abnormal behavior recognition for intelligent video surveillance systems: A review. *Expert Systems with Applications*.

Machado I. P., Gomes, A. L., Gamboa, H., Paixão, V., & Costa, R. M. (2015). Human activity data discovery from triaxial accelerometer sensor: Non-supervised learning sensitivity to feature extraction parametrization. *Information Processing & Management, 51*, 204-214.

Mannini A., Intille, S. S., Rosenberger, M., Sabatini, A. M., & Haskell, W. (2013). Activity recognition using a single accelerometer placed at the wrist or ankle. *Medicine and science in sports and exercise, 45*, 2193.

Mannini A., & Sabatini, A. M. (2010). Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors, 10*, 1154-1175.

Mannini A., & Sabatini, A. M. (2011). Accelerometry-based classification of human activities using Markov modeling. *Computational intelligence and neuroscience, 2011*, 4.

Margarito J., Helaoui, R., Bianchi, A. M., Sartor, F., & Bonomi, A. G. (2016). User-independent recognition of sports activities from a single wrist-worn accelerometer: A template-matching-based approach. *IEEE Transactions on Biomedical Engineering, 63*, 788-796.

Masse F., Gonzenbach, R., Paraschiv-Ionescu, A., Luft, A. R., & Aminian, K. (2016). Wearable barometric pressure sensor to improve postural transition recognition of mobility-impaired stroke patients. *IEEE transactions on neural systems and rehabilitation engineering, 24*, 1210-1217.

Maurer U., Smailagic, A., Siewiorek, D. P., & Deisher, M. (2006). Activity recognition and monitoring using multiple sensors on different body positions. In *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on* (pp. 4 pp.-116): IEEE.

Mehr H. D., Polat, H., & Cetin, A. (2016). Resident activity recognition in smart homes by using artificial neural networks. In *Smart Grid Congress and Fair (ICSG), 2016 4th International Istanbul* (pp. 1-5): IEEE.

Mehrang S., Pietila, J., Tolonen, J., Helander, E., Jimison, H., Pavel, M., & Korhonen, I. (2017). Human Activity Recognition Using A Single Optical Heart Rate Monitoring Wristband Equipped with Triaxial Accelerometer. In *EMBEC & NBC 2017* (pp. 587-590): Springer.

Meyer P. E., & Bontempi, G. (2006). On the use of variable complementarity for feature selection in cancer classification. In *Workshops on Applications of Evolutionary Computation* (pp. 91-102): Springer.

Mimobaby. (2018). Sleep trackers for little ones. In (Vol. 2018).

Mohamed A. B. H., Val, T., Andrieux, L., & Kachouri, A. (2012). Using a Kinect WSN for home monitoring: principle, network and application evaluation. In *Wireless Communications in Unusual and Confined Areas (ICWCUCA), 2012 International Conference on* (pp. 1-5): IEEE.

Moncada-Torres A., Leuenberger, K., Gonzenbach, R., Luft, A., & Gassert, R. (2014). Activity classification based on inertial and barometric pressure sensors at different anatomical locations. *Physiological measurement, 35*, 1245.

Montalto F., Guerra, C., Bianchi, V., De Munari, I., & Ciampolini, P. (2015). MuSA: Wearable Multi Sensor Assistant for Human Activity Recognition and Indoor Localization. In *Ambient Assisted Living* (pp. 81-92): Springer.

Morales J., & Akopian, D. (2017). Physical activity recognition by smartphones, a survey. *Biocybernetics and Biomedical Engineering, 37*, 388-400.

Morales J., Akopian, D., & Agaian, S. (2014). Human activity recognition by smartphones regardless of device orientation. In *IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics* (pp. 90300I-90300I).

Mortazavi B., Nemati, E., VanderWall, K., Flores-Rodriguez, H. G., Cai, J. Y. J., Lucier, J., Naeim, A., & Sarrafzadeh, M. (2015). Can smartwatches replace smartphones for posture tracking? *Sensors, 15*, 26783-26800.

Mortazavi B. J., Pourhomayoun, M., Alsheikh, G., Alshurafa, N., Lee, S. I., & Sarrafzadeh, M. (2014). Determining the single best axis for exercise repetition recognition and counting on smartwatches. In *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on* (pp. 33-38): IEEE.

Murao K., & Terada, T. (2014). A recognition method for combined activities with accelerometers. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (pp. 787-796): ACM.

Nakamura M., Nakamura, J., Shuzo, M., Warisawa, S., & Yamada, I. (2010). Collaborative processing of wearable and ambient sensor system for health monitoring application. In *Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on* (pp. 1-5): IEEE.

Nam Y., & Park, J. W. (2013). Child activity recognition based on cooperative fusion model of a triaxial accelerometer and a barometric pressure sensor. *IEEE journal of biomedical and health informatics, 17*, 420-426.

Ngiam J., Chen, Z., Bhaskar, S. A., Koh, P. W., & Ng, A. Y. (2011). Sparse filtering. In *Advances in neural information processing systems* (pp. 1125-1133).

Nweke H. F., Teh, Y. W., Al-Garadi, M. A., & Alo, U. R. (2018). Deep Learning Algorithms for Human Activity Recognition using Mobile and Wearable Sensor Networks: State of the Art and Research Challenges. *Expert Systems with Applications*.

Ordóñez F. J., Iglesias, J. A., De Toledo, P., Ledezma, A., & Sanchis, A. (2013). Online activity recognition using evolving classifiers. *Expert Systems with Applications, 40*, 1248-1255.

Pang I., Okubo, Y., Sturnieks, D., Lord, S. R., & Brodie, M. A. (2019). Detection of near falls using wearable devices: a systematic review. *Journal of geriatric physical therapy, 42*, 48-56.

Pansiot J., Stoyanov, D., McIlwraith, D., Lo, B., & Yang, G. (2007). Ambient and wearable sensor fusion for activity recognition in healthcare monitoring systems. In *4th international workshop on wearable and implantable body sensor networks (BSN 2007)* (pp. 208-212): Springer.

Panwar M., Dyuthi, S. R., Prakash, K. C., Biswas, D., Acharyya, A., Maharatna, K., Gautam, A., & Naik, G. R. (2017). 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* (pp. 2438-2441): IEEE.

Parkka J., Ermes, M., Korpipaa, P., Mantyjarvi, J., Peltola, J., & Korhonen, I. (2006). Activity classification using realistic data from wearable sensors. *IEEE Transactions on information technology in biomedicine, 10*, 119-128.

Patil T. R., & Sherekar, S. (2013). Performance analysis of Naive Bayes and J48 classification algorithm for data classification. *International Journal of Computer Science and Applications, 6*, 256-261.

Pavey T. G., Gilson, N. D., Gomersall, S. R., Clark, B., & Trost, S. G. (2017). Field evaluation of a random forest activity classifier for wrist-worn accelerometer data. *Journal of science and medicine in sport, 20*, 75-80.

Peng H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence, 27*, 1226-1238.

Peng L., Chen, L., Wu, X., Guo, H., & Chen, G. (2017). Hierarchical complex activity representation and recognition using topic model and classifier level fusion. *IEEE Transactions on Biomedical Engineering, 64*, 1369-1379.

Phillips L. J., DeRoche, C. B., Rantz, M., Alexander, G. L., Skubic, M., Despins, L., Abbott, C., Harris, B. H., Galambos, C., & Koopman, R. J. (2017). Using embedded sensors in independent living to predict gait changes and falls. *Western journal of nursing research, 39*, 78-94.

Plötz T., Hammerla, N. Y., & Olivier, P. (2011). Feature learning for activity recognition in ubiquitous computing. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence* (Vol. 22, pp. 1729).

Pourbabaee B., Roshtkhari, M. J., & Khorasani, K. (2017). Deep Convolution Neural Networks and Learning ECG Features for Screening Paroxysmal Atrial Fibrillatio Patients. *IEEE Transactions on Systems, Man, and Cybernetics: Systems.*

Raja K. B., Raghavendra, R., Vemuri, V. K., & Busch, C. (2015). Smartphone based visible iris recognition using deep sparse filtering. *Pattern Recognition Letters, 57*, 33-42.

Ramos-Garcia R. I., & Hoover, A. W. (2013). A study of temporal action sequencing during consumption of a meal. In *Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics* (pp. 68): ACM.

Rasheed M. B., Javaid, N., Alghamdi, T. A., Mukhtar, S., Qasim, U., Khan, Z. A., & Raja, M. H. B. (2015). Evaluation of human activity recognition and fall detection using android phone. In *Advanced Information Networking and Applications (AINA), 2015 IEEE 29th International Conference on* (pp. 163-170): IEEE.

Ravi D., Wong, C., Lo, B., & Yang, G.-Z. (2016). 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* (pp. 71-76): IEEE.

Ravi N., Dandekar, N., Mysore, P., & Littman, M. L. (2005). Activity recognition from accelerometer data. In *Aaai* (Vol. 5, pp. 1541-1546).

Reddy S., Mun, M., Burke, J., Estrin, D., Hansen, M., & Srivastava, M. (2010). Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks (TOSN), 6*, 13.

Reiss A., Hendeby, G., & Stricker, D. (2015). A novel confidence-based multiclass boosting algorithm for mobile physical activity monitoring. *Personal and Ubiquitous Computing, 19*, 105-121.

Reiss A., & Stricker, D. (2012). Introducing a new benchmarked dataset for activity monitoring. In *Wearable Computers (ISWC), 2012 16th International Symposium on* (pp. 108-109): IEEE.

Reyes-Ortiz J.-L., Oneto, L., Sama, A., Parra, X., & Anguita, D. (2016). Transition-aware human activity recognition using smartphones. *Neurocomputing, 171*, 754-767.

Rodriguez-Martin D., Samà, A., Perez-Lopez, C., Català, A., Cabestany, J., & Rodriguez-Molinero, A. (2013). SVM-based posture identification with a single waist-located triaxial accelerometer. *Expert Systems with Applications, 40*, 7203-7211.

Ronao C. A., & Cho, S.-B. (2015). Deep convolutional neural networks for human activity recognition with smartphone sensors. In *International Conference on Neural Information Processing* (pp. 46-53): Springer.

Ronao C. A., & Cho, S.-B. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Systems with Applications, 59*, 235-244.

Roy N., Misra, A., & Cook, D. (2016). Ambient and smartphone sensor assisted ADL recognition in multi-inhabitant smart environments. *Journal of ambient intelligence and humanized computing, 7*, 1-19.

Sani S., Massie, S., Wiratunga, N., & Cooper, K. (2017). Learning Deep and Shallow Features for Human Activity Recognition. In *International Conference on Knowledge Science, Engineering and Management* (pp. 469-482): Springer.

Sathyanarayana A., Joty, S., Fernandez-Luque, L., Ofli, F., Srivastava, J., Elmagarmid, A., Arora, T., & Taheri, S. (2016). Sleep quality prediction from wearable data using deep learning. *JMIR mHealth and uHealth, 4*.

Schölkopf B., Smola, A., & Müller, K.-R. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation, 10*, 1299-1319.

Shibuya N., Nukala, B. T., Rodriguez, A., Tsay, J., Nguyen, T. Q., Zupancic, S., & Lie, D. Y. (2015). A real-time fall detection system using a wearable gait analysis sensor and a support vector machine (svm) classifier. *Mobile Computing and Ubiquitous Networking (ICMU), 2015 Eighth International Conference on*, 66-67.

Shinmura S. (2014). Comparison of Linear Discriminant Function by K-fold Cross-validation. *Data Analytic, 2014*, 1-6.

Shoaib M., Bosch, S., Incel, O., Scholten, H., & Havinga, P. (2016). Complex human activity recognition using smartphone and wrist-worn motion sensors. *Sensors, 16*, 426.

Sivakumar P., & Shajina, T. (2012). Human gait recognition and classification using time series shapelets. In *IEEE international conference on advances in computing and communications* (Vol. 33).

Song S., Chandrasekhar, V., Mandal, B., Li, L., Lim, J.-H., Sateesh Babu, G., Phyo San, P., & Cheung, N.-M. (2016). Multimodal multi-stream deep learning for egocentric activity recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 24-31).

Stikic M., Huynh, T., Van Laerhoven, K., & Schiele, B. (2008). ADL recognition based on the combination of RFID and accelerometer sensing. *Pervasive Computing Technologies for Healthcare, 2008. PervasiveHealth 2008. Second International Conference on*, 258-263.

Stikic M., Van Laerhoven, K., & Schiele, B. (2008). Exploring semi-supervised and active learning for activity recognition. *Wearable computers, 2008. ISWC 2008. 12th IEEE international symposium on*, 81-88.

Stone E. E., & Skubic, M. (2015). Fall detection in homes of older adults using the Microsoft Kinect. *IEEE journal of biomedical and health informatics, 19*, 290-301.

Su X., Tong, H., & Ji, P. (2014). Activity recognition with smartphone sensors. *Tsinghua Science and Technology, 19*, 235-249.

Subrahmanya N., & Shin, Y. C. (2010). Sparse multiple kernel learning for signal processing applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*, 788-798.

Sun L., Zhang, D., Li, B., Guo, B., & Li, S. (2010). Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations. *Ubiquitous intelligence and computing*, 548-562.

Suto J., Oniga, S., Lung, C., & Orha, I. (2017). Recognition rate difference between real-time and offline human activity recognition. In *Internet of Things for the Global Community (IoTGC), 2017 International Conference on* (pp. 1-6): IEEE.

Suto J., Oniga, S., & Sitar, P. P. (2016). Feature analysis to human activity recognition. *International Journal of Computers Communications & Control, 12*, 116-130.

Sztyler T., Stuckenschmidt, H., & Petrich, W. (2017). Position-aware activity recognition with wearable devices. *Pervasive and Mobile Computing, 38*, 281-295.

Tapia E. M., Intille, S. S., Haskell, W., Larson, K., Wright, J., King, A., & Friedman, R. (2007). Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In *Wearable Computers, 2007 11th IEEE International Symposium on* (pp. 37-40): IEEE.

Terada T., & Tanaka, K. (2010). A framework for constructing entertainment contents using flash and wearable sensors. In *International Conference on Entertainment Computing* (pp. 334-341): Springer.

Tunca C., Alemdar, H., Ertan, H., Incel, O. D., & Ersoy, C. (2014). Multimodal wireless sensor network-based ambient assisted living in real homes with multiple residents. *Sensors, 14*, 9692-9719.

Uddin M., Salem, A., Nam, I., & Nadeem, T. (2015). Wearable sensing framework for human activity monitoring. In *Proceedings of the 2015 workshop on Wearable Systems and Applications* (pp. 21-26): ACM.

Um T. T., Babakeshizadeh, V., & Kulic, D. (2016). Exercise Motion Classification from Large-Scale Wearable Sensor Data Using Convolutional Neural Networks. *arXiv preprint arXiv:1610.07031*.

Um T. T., Pfister, F. M., Pichler, D., Endo, S., Lang, M., Hirche, S., Fietzek, U., & Kulić, D. (2017). Data augmentation of wearable sensor data for parkinson's disease monitoring using convolutional neural networks. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction* (pp. 216-220): ACM.

United Nations. (2017). World Population Prospects:Key findings & advance tables. In (Vol. 2017 Revision).

Uray M., Skocaj, D., Roth, P. M., Bischof, H., & Leonardis, A. (2007). Incremental LDA Learning by Combining Reconstructive and Discriminative Approaches. In *BMVC* (Vol. 2007, pp. 272-281).

Uslu G., Dursunoglu, H. I., Altun, O., & Baydere, S. (2013). Human activity monitoring with wearable sensors and hybrid classifiers. *International Journal of Computer Information Systems and Industrial Management Applications, 5*, 345-353.

Van Der Maaten L., Postma, E., & Van den Herik, J. (2009). Dimensionality reduction: a comparative. *J Mach Learn Res, 10*, 66-71.

van Kasteren Y., Bradford, D., Zhang, Q., Karunanithi, M., & Ding, H. (2017). Understanding Smart Home Sensor Data for Ageing in Place Through Everyday Household Routines: A Mixed Method Case Study. *JMIR mHealth and uHealth, 5*, e52.

Vavoulas G., Chatzaki, C., Malliotakis, T., Pediaditis, M., & Tsiknakis, M. (2016). The MobiAct Dataset: Recognition of Activities of Daily Living using Smartphones. In *ICT4AgeingWell* (pp. 143-151).

Vehtari A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing, 27*, 1413-1432.

Vepakomma P., De, D., Das, S. K., & Bhansali, S. (2015). A-Wristocracy: Deep learning on wrist-worn sensing for recognition of user complex activities. In *Wearable and Implantable Body Sensor Networks (BSN), 2015 IEEE 12th International Conference on* (pp. 1-6): IEEE.

Wang A., Chen, G., Yang, J., Zhao, S., & Chang, C.-Y. (2016). A comparative study on human activity recognition using inertial sensors in a smartphone. *IEEE Sensors Journal, 16*, 4566-4578.

Wang J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2017). Deep learning for sensor-based activity recognition: A survey. *arXiv preprint arXiv:1707.03502*.

Wang J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2018). Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*.

Wang J., Zhang, X., Gao, Q., Yue, H., & Wang, H. (2017). Device-free wireless localization and activity recognition: A deep learning approach. *IEEE Transactions on Vehicular Technology, 66*, 6258-6267.

Wang L. (2016). Recognition of human activities using continuous autoencoders with wearable sensors. *Sensors, 16*, 189.

Wang L., Gu, T., Xie, H., Tao, X., Lu, J., & Huang, Y. (2013). A wearable rfid system for real-time activity recognition using radio patterns. In *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services* (pp. 370-383): Springer.

Wang S., Yang, J., Chen, N., Chen, X., & Zhang, Q. (2005). Human activity recognition with user-free accelerometers in the sensor networks. In *Neural Networks and Brain, 2005. ICNN&B'05. International Conference on* (Vol. 2, pp. 1212-1217): IEEE.

Wang Y., Cang, S., & Yu, H. (2016). Realization of wearable sensors-based human activity recognition with an augmented feature group. In *Automation and Computing (ICAC), 2016 22nd International Conference on* (pp. 473-478): IEEE.

Wang Y., Cang, S., & Yu, H. (2018). A Data Fusion based Hybrid Sensory System for Older People's Daily Activity and Daily Routine Recognition. *IEEE Sensors Journal*.

Wang Z., Yang, Z., & Dong, T. (2017). A review of wearable technologies for elderly care that can accurately track indoor position, recognize physical activities and monitor vital signs in real time. *Sensors, 17*, 341.

Wilson D. H., & Atkeson, C. (2005). Simultaneous tracking and activity recognition (STAR) using many anonymous, binary sensors. In *International Conference on Pervasive Computing* (pp. 62-79): Springer.

Wu J., Wang, J., & Liu, L. (2007). Feature extraction via KPCA for classification of gait patterns. *Human movement science, 26*, 393-411.

Wu W., Dasgupta, S., Ramirez, E. E., Peterson, C., & Norman, G. J. (2012). Classification accuracies of physical activities using smartphone motion sensors. *Journal of medical Internet research, 14*.

Xi X., Tang, M., Miran, S. M., & Luo, Z. (2017). Evaluation of feature extraction and recognition for activity monitoring and fall detection based on wearable sEMG sensors. *Sensors, 17*, 1229.

Xing Z., Pei, J., Yu, P. S., & Wang, K. (2011). Extracting interpretable features for early classification on time series. In *Proceedings of the 2011 SIAM International Conference on Data Mining* (pp. 247-258): SIAM.

Yang M., Zheng, H., Wang, H., McClean, S., Hall, J., & Harris, N. (2012). A machine learning approach to assessing gait patterns for complex regional pain syndrome. *Medical Engineering and Physics, 34*, 740-746.

Yao S., Hu, S., Zhao, Y., Zhang, A., & Abdelzaher, T. (2017). Deepsense: A unified deep learning framework for time-series mobile sensing data processing. In *Proceedings of the 26th International Conference on World Wide Web* (pp. 351-360): International World Wide Web Conferences Steering Committee.

Yoon S., Sim, J. K., & Cho, Y.-H. (2016). A flexible and wearable human stress monitoring patch. *Scientific reports, 6*, 23468.

Yu L., Xiong, D., Guo, L., & Wang, J. (2016). A compressed sensing-based wearable sensor network for quantitative assessment of stroke patients. *Sensors, 16*, 202.

Zakaria J., Mueen, A., Keogh, E., & Young, N. (2016). Accelerating the discovery of unsupervised-shapelets. *Data mining and knowledge discovery, 30*, 243-281.

Zeng M., Nguyen, L. T., Yu, B., Mengshoel, O. J., Zhu, J., Wu, P., & Zhang, J. (2014). Convolutional neural networks for human activity recognition using mobile sensors. In *Mobile Computing, Applications and Services (MobiCASE), 2014 6th International Conference on* (pp. 197-205): IEEE.

Zhan K., Ramos, F., & Faux, S. (2012). Activity recognition from a wearable camera. In *Control Automation Robotics & Vision (ICARCV), 2012 12th International Conference on* (pp. 365-370): IEEE.

Zhang J., & Wu, Y. (2018). Automatic sleep stage classification of single-channel EEG by using complex-valued convolutional neural network. *Biomedical Engineering/Biomedizinische Technik, 63*, 177-190.

Zhang M., & Sawchuk, A. A. (2011). A feature selection-based framework for human activity recognition using wearable multimodal sensors. In *Proceedings of the 6th International Conference on Body Area Networks* (pp. 92-98): ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

Zhang M., & Sawchuk, A. A. (2012). Motion primitive-based human activity recognition using a bag-of-features approach. In *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium* (pp. 631-640): ACM.

Zhang S., McCullagh, P., & Callaghan, V. (2014). An efficient feature selection method for activity classification. In *Intelligent Environments (IE), 2014 International Conference on* (pp. 16-22): IEEE.

Zhang S., Wei, Z., Nie, J., Huang, L., Wang, S., & Li, Z. (2017). A Review on Human Activity Recognition Using Vision-Based Method. *Journal of healthcare engineering, 2017*.

Zheng Y., Liu, Q., Chen, E., Ge, Y., & Zhao, J. L. (2014). Time series classification using multi-channels deep convolutional neural networks. In *International Conference on Web-Age Information Management* (pp. 298-310): Springer.

Zheng Y., Wong, W.-K., Guan, X., & Trost, S. (2013). Physical Activity Recognition from Accelerometer Data Using a Multi-Scale Ensemble Method. In *IAAI*.

Zhou B., Sundholm, M., Cheng, J., Cruz, H., & Lukowicz, P. (2016). Measuring muscle activities during gym exercises with textile pressure mapping sensors. *Pervasive and Mobile Computing*.

Zolfaghari S., & Keyvanpour, M. R. (2016). SARF: Smart activity recognition framework in Ambient Assisted Living. In *Computer Science and Information Systems (FedCSIS), 2016 Federated Conference on* (pp. 1435-1443): IEEE.