A Comprehensive Survey on Human Activity Recognition Using Sensing Technology

Ming Wang

College of Information Engineering, Chang'an University, Xi'an, China.

2363659293@qq.com

Abstract. With the rapid development of machine learning technology and the popularity of various electronic devices such as smartphones and smart bracelets, sensor technology has been applied to all aspects of human society. On this basis, HAR already plays an important role in various fields including sports and the military. In the medical field, particularly, HAR is gaining increasing recognitions by researchers to be one of the main technologies for future behavioral detection in the elderly. The application of HAR technology allows users to identify human behavior through applications equipped with specific sensors, reducing human costs. This survey sorted out the published articles related to HAR research in recent years, provided an analysis of the background, applications, and advantages and disadvantages of sensing technologies, and made a brief exposition of the HAR process.

Keywords: Human Activity Recognition (HAR), machine learning (ML), deep learning (DL), sensors, Activity Recognition.

1. Introduction

In the current age with rapid renewal of microelectronics and computer technology, sensor technology has been applied into wider scope, greatly influencing human life and habits. Sensors have many advantages in applications such as the detection of various information such as temperature and humidity; on the other hand, sensor technology can also be applied in devices that humans can often access, such as smartphones [1], smart homes devices, etc. One representative product is wearable sensors, which are particularly important in the field of healthcare. For example, wearable sensors make it possible to monitor the daily activities of patients suffering from cancer, dementia, Alzheimer's disease, stroke and other diseases, helping patients who are unable to take care of themselves to live with minimal assistance from their families and doctors.

Although human activity recognition (HAR) technology has been implemented in a number of areas, it still faces a number of challenges, including (i) limitations in the contact surface carried by wearable devices, (ii) sensor accuracy limitations, (iii) cost limitations, (iv) privacy data access and protection issues, (v) difficulties in statistics and interpretation of the acquired data, (vi) difficulties in accurately quantifying the efficiency of HAR systems, (vii) the selection of HAR algorithm models problem, (viii) variability among different individuals leading to the need for retraining of new users, and (ix) other contingencies in diverse and complex real-world environments [2]. With the development of sensing technology, HAR technology will also become more mature and these problems will gradually be solved effectively. Therefore, in order to identify human activities more effectively, it is crucial to investigate how to acquire data efficiently.

The HAR system relies on three main ways of acquiring data: (i) Data can be obtained through environmental sensors, fixing temperature, humidity and other sensors on specific objects, such as smart homes, floors, cameras, taps, WiFi devices, etc. Information about the surrounding environment is collected from the surroundings, the information is processed and converted into relevant data for analysis. (ii) Data can be obtained through wearable sensors. Tiny sensors such as accelerometer sensors [3], gyroscopes, GPS sensors, compasses, heart rate sensors, etc. are embedded in mobile phones, smart bracelets, clothing and then worn around the tester's ears, around the neck, around the waist, at the legs and feet or inside the pockets, following the wearer as they move, thus allowing for longer and more comprehensive access to information about the wearer. (iii) Data can also be obtained

through a mix of sensors. For example, environmental sensors can be combined with smartphones to create a multi-resident smart environment [4]. This can be done by fixing the environmental sensor in a hidden environment for the individual, and then having the test subject move around the environment with a smartphone or other wearable sensor, while acquiring data from the micro-sensor integrated into the wearable sensor and the environmental sensor. The data obtained in both ways are compared to each other and complement each other to compensate for the shortcomings of the environmental sensor and the wearable sensor and to obtain more comprehensive sensing information.

The comparison of the three ways of acquiring data indicates that none of them are perfect in all sense. The advantages of HAR systems based on environmental sensors are obvious. The data acquired is more complete in a limited space, and there are often no size or weight limitations, allowing for more and more complete sensors to be integrated into the system. The disadvantages are also clear, with space constraints and privacy issues posing a significant resistance to the development of such HARs. Wearable sensors are more versatile in practical applications and do not have space constraints, allowing access to more comprehensive information about the subject, but are limited by weight and size, and the advantages of such HAR systems will become more apparent with the future development of highly sophisticated sensors. Hybrid sensor-based HAR systems integrate the advantages of both sensors and can obtain more comprehensive information, but are more costly and currently have few combined solutions, which will have considerable potential in the future. Table 1 shows specific descriptions and paper citations of the three ways in which HAR systems can acquire data, and compares their advantages and disadvantages. The statistical content of the table shows that HAR systems have a wide range of application areas and that the application of data acquisition via wearable sensors in particular is flexible and will have a very prominent position in future HAR research. HAR systems have two options in terms of specific algorithm implementation, one being traditional machine learning algorithms and the other deep learning algorithms.

Table 1 Data Acquisition Methods of HAR Systems

Access to data	Description	Related Sensors	Equipment	Advantages	Disadvantages	Thesis Citation
Environmental Sensors	Fixing the sensor to a specific object or environment to obtain data	Temperature sensors, humidity sensors, passive infrared sensors, etc.	Surveillance cameras, smart homes, flooring, doors, WIFI devices, etc.	Complete data acquired in a limited space, generally without volume or weight restrictions	Restricted space, risk of privacy invasion, expensive	[12], [15], [22], [23], [30], [39], [46], [48]
Wearable Sensors	Integration of sensors into wearable devices to capture data	Accelerometer sensors, gyroscopes, GPS sensors, compass, heart rate sensors, etc.	Smartphones, smartwatches, etc.	High flexibility, unrestricted space, small size and low cost	Low wearer comfort, high precision component design, high production difficulty	[1], [3], [9], [14], [16], [17], [19], [20], [38], [45]

Hybrid Sensors	Combination of environmental and wearable sensors for joint data acquisition	Temperature sensors, humidity sensors, accelerometer sensors, gyroscopes, etc.	Integration of surveillance cameras, smart homes, WIFI devices, etc. with smartphones, smart watches, etc.	The advantages of multiple sensors for better information acquisition	High complexity and high cost	[4], [10], [38]

It is hard to judge which algorithm performs better since they excel in different perspectives, which suits for different demands of a certain field.

This paper studies recent advances in HAR research, providing a more comprehensive account of the components, applications, and algorithms of HAR systems and other information. In the survey, we followed standard steps to analyze HAR research. First, we collected papers on HAR research, and then, the collected articles were screened and analyzed.

In surveying the articles, we have proposed the following main research questions.

What questions relevant to HAR research are addressed in these articles?

1. Which algorithm is used?
2. Which sensors are used to obtain data?
3. Which equipment is it used in?
4. What is the accuracy rate of the algorithm?

The rest of the article is organized as follows: Section 2 introduces background knowledge related to machine learning, deep learning, and sensing technologies. Section 3 describes HAR applications in healthcare and sports. Section 4 describes the HAR process, including sensor selection, data pre-processing, feature extraction, dimensionality reduction, classification and recognition, and Section 5 summarizes the HAR survey.

2. Related Work

In practical applications, HAR systems begin by sensing human activity through sensors such as temperature and accelerometers and converting them into electrical signals, and processing the information collected by the sensors through traditional machine learning or deep learning and other algorithms to recognize human activity. This section will provide a brief overview of the background related to HAR research, including machine learning, deep learning and sensing technologies.

2.1 Machine Learning

With the advent of the Big Data era, machine learning has been a hot topic and more and more people believe that it will become an essential discipline in the world of tomorrow. When it comes to machine learning, there is no getting around artificial intelligence and computer science.

Artificial intelligence is subordinate to computer science and is a very important discipline in computer science that aims to simulate the human brain and behavior through observation, learning and research so that computer systems or other machines can mimic the way humans think or act and move, thus extending human organs such as the five senses and limbs so that humans can solve problems more quickly and efficiently.

Machine learning is subordinate to artificial intelligence and is an important discipline in artificial intelligence, which largely covers a number of disciplines such as statistics and computer science, and is a mature discipline in terms of both theory and application. In short, machine learning is a discipline that studies algorithms for the analysis and optimization of input information through data or past experience [5].

When studying machine learning, it can be analyzed at the level of learning methods and algorithms. In terms of learning methods, machine learning algorithms can be mainly divided into supervised learning and unsupervised learning.

2.1.1 Supervised learning:

Also known as tutored learning, supervised learning is a learning task that learns from a pre-determined and labeled training dataset and generates a function [6]. The functions generated by this learning task have predictive capabilities, so that when a human input a test dataset to a machine, the machine can use the learned functions to predict the outcome and solve a classification or regression problem. Figure 1 shows the general flow chart of supervised learning.

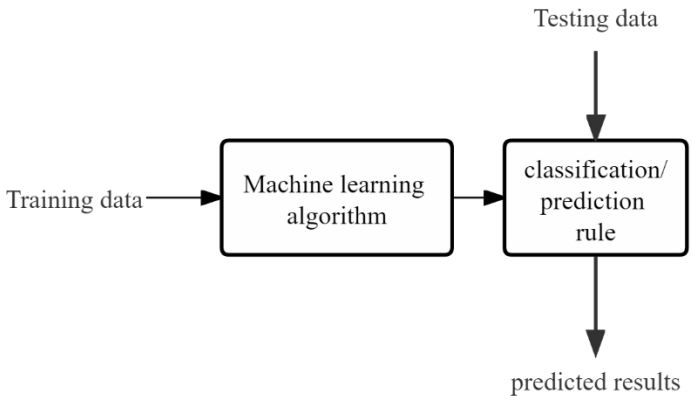


Figure 1 General Flow Chart of Supervised Learning

Supervised learning is often used to solve regression and classification problems. Predictions for regression problems are often continuous and specific values, such as predicting house prices, temperatures, credit values, etc. Predictions for classification problems are often discrete values, for example, identifying the species and gender of organisms in a picture. When using supervised learning algorithms to solve regression problems, the algorithms often used are linear regression, regression trees, K-Nearest Neighbours(KNN), Adaboosting and neural networks. The first two of these algorithms can only be used to solve regression problems, while the last three can be used to solve both regression and classification problems. When solving classification problems, algorithms such as plain Bayes, decision trees, support vector machines (SVM) and logistic regression are often used.

2.1.2 Unsupervised learning:

*A*lso known as tutorless learning, unsupervised learning is another type of machine learning approach that is distinguished from supervised learning in that it does not have a clear objective and the data it uses is unlabelled, making it impossible to judge whether the test results are sound or not[7]. This type of learning is suitable for problems that lack a priori knowledge but require classification of things, such as discovering unusual information in a pile of data, classifying users based on their identity, age and other portraits, etc.

In terms of algorithms, unsupervised learning is more widely used in two algorithms, namely clustering and dimensionality reduction. Clustering algorithms include K-means clustering and hierarchical clustering algorithms, while dimensionality reduction algorithms include principal component analysis (PCA) and singular value decomposition (SVD) algorithms. Clustering has similarities to classification, but is completely different. When solving classification problems in supervised learning, the researcher knows in advance which categories the test results will be and which categories are correct and which are incorrect, whereas in clustering systems, the categories of test results are more abstract and difficult to describe concretely, and the researcher does not know which categories the test results will be and exactly what they mean. A dimensionality reduction

algorithm is similar to compression, an algorithm that maintains the structure of the data to the greatest extent possible and compresses the dimensionality of the data.

2.2 Deep Learning

Deep learning, which is subordinate to machine learning, is the most important and hotly researched area of machine learning, and with the emergence of deep learning, there has been a third high point in the development of machine learning, which has received more attention [8]. Deep learning is stripped down from traditional neural networks and is a method that follows the optimization and enrichment of traditional neural networks. Deep learning systems rely on human-designed computer computing systems and the training and learning of large amounts of data, imitating the connections between neurons in the human brain and computing neural signals. The goal of deep learning is to be able to highly imitate the human brain's thinking and to achieve efficient recognition and processing of data such as text, speech and images with the advantage of the computer's high-speed computing power.

Compared with traditional machine learning algorithms, the biggest difference of deep learning algorithms from others is that the features are extracted by the system itself, without relying on human operations, which saves human time to a large extent. Although human extraction of features is not necessarily time-consuming and labor-intensive, and sometimes even faster and more accurate than system extraction, there are differences between models, and sometimes the differences can be significant, increasing the difficulty of human extraction, so deep learning has a great advantage in feature extraction. In contrast, the disadvantages of deep learning are also obvious, as features are extracted by machines and it is difficult to understand the exact extraction principles in practice.

There are many advantages to deep learning beyond the extraction of features. Compared to traditional machine learning, deep learning has a much better learning instinct, especially in the areas of speech and image recognition, where it has shown excellent learning capabilities and great potential. Also, because of the high demand for data training and the low demand for human involvement in the models, deep learning currently has a higher and more far-reaching ceiling than traditional machine learning can foresee. At present, deep learning is already doing better than humans in some aspects of imagery, facial recognition and other areas.

Deep learning now also has a number of drawbacks. On the one hand, deep learning is computationally intensive and difficult to keep costs low as many deep learning systems are difficult to integrate into mobile applications and require high performance GPUs to train data in a short period of time with high computation. On the other hand, when the data set is small, deep learning has no advantage over traditional machine learning, and in many cases is no match for it.

In terms of algorithms, there are four more commonly used algorithms for deep learning, namely Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Generative Adversarial Networks (GAN) and Deep Reinforcement Learning (RL) algorithms.

2.3 Sensing Technology

Sensors are often essential in the study of human activity recognition, as information relating to walking, running etc. must first be collected for the study of human activity, and then processed, analyzed and evaluated. A sensor is a device that senses static or dynamic information about its surroundings, such as temperature and humidity, through sensitive elements, converts this information into an electrical or other signal through conversion elements and transmits this signal to other electronic devices such as processors. Figure 2 illustrates the main components of a general sensor.

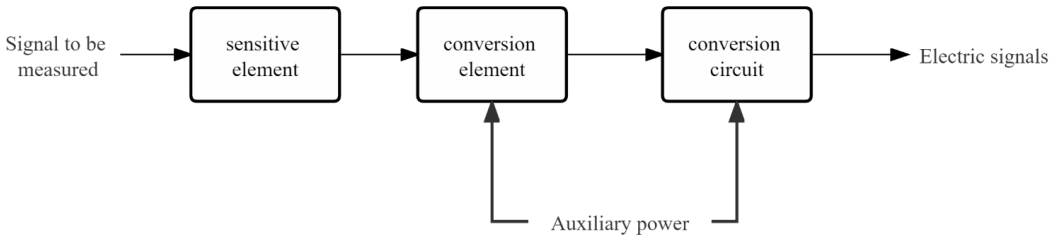


Figure 2 Main Components of a General Sensor

Among the sensors that detect daily human activity, accelerometer sensors have a greater advantage. Accelerometer sensors can more accurately identify the exact numerical magnitude of the acceleration of a measured target in the x, y and z axes and thus determine how the measured target is moving horizontally and vertically in space. By integrating accelerometer sensors into wearable devices and wearing them on the chest, back, wrists, legs, waist, feet and other parts of the body, complex activity information such as walking, scrubbing, drinking and eating can be sensed.

Sensor fusion technology will be a major direction in the development of sensors in the future, as opposed to through a single sensor. Sensor fusion technology is a technology that integrates the information acquired by multiple sensors through computer technology to analyze and process them. For example, the fusion of accelerometers and magnetometers can provide a more comprehensive sense of the position movement of the object under test in slow motion [9].

In many cases, the data sensed by the sensors should be multifaceted, just as the human brain processes information through touch, vision, hearing and other multi-dimensional senses. The data obtained from multiple sensors will have a higher accuracy [10]. In addition, the advantages of fused sensors are clear and huge, with the development of sensor technology, the size of various sensors will continue to reduce, fusion sensor size, bulky and other defects will be effectively made up, which will make sensor fusion technology in the future to get more applications.

In terms of algorithms, there are two main algorithms for sensor fusion technology, the first being stochastic methods such as weighted averaging, Kalman filtering, multi-Bayes estimation, D-S evidential inference and generative rules, and the other being artificial intelligence methods such as fuzzy logic inference and artificial neural network methods.

In terms of applications, one of the more common applications of sensor fusion technology is robot perception. For robots, sensing the environment is a more complex and difficult capability, as often robots cannot accurately determine their surroundings through information such as images or sounds alone. In robot haptics alone, the robot needs to be able to recognize various information such as the temperature, size, shape and material of the object it is touching, which requires the use of a variety of sensors such as temperature and pressure to analyze the category, size and state of the object it is touching through a combination of various information [11].

3. Fields HAR Applied

In the past two decades, human activity recognition has developed well in the fields of healthcare, sports, and physical games. The rapid development of the sensor industry has led to a trend towards sophisticated sensor functions that can be kept tiny and embedded in various devices, thus making the future of HAR research even more promising [12]. In this section, the applications of HAR research are described in the fields of healthcare and sports respectively.

3.1 Healthcare

The aging of population has become the trend for the whole society in the next 30 years, and by 2025 the number of people aged 60 and over will exceed 20% of the human population [13].

In traditional healthcare, patients are often housed in wards and cared for by their relatives or medical staff, and human care often involves uncertainty, such as abnormalities in the patient's health during the caregiver's rest time, which can make it difficult to detect abnormalities in time and miss the best treating time. To prevent this from happening, patients can be equipped with wearable devices equipped with wireless sensor networks that can sense and analyze physical characteristics such as ECG and blood oxygen information over a long period of time in the patient's home [14]. The development from ward-centric medical monitoring systems to patient-centric medical monitoring systems is a great advancement in the healthcare industry, which will effectively reduce healthcare and manpower costs, and enable the first detection of physical abnormalities in patients and then contact medical personnel for assistance.

Other typical applications of HAR research in the medical field include the use of passive WIFI sensing to identify human respiratory activity and analyze breathing rates to determine the health status of patients [15], and the analysis of brain waves during sleep to determine sleep quality. It can also monitor the daily activities of orphans and elderly people, and send medical advice to patients in case of dangerous accidents such as falls, and send danger notifications to patients' families or medical staff [16]. Especially for elderly people with diseases such as Alzheimer's and epilepsy, where long-term hospitalisation is not practical, the HAR system can be used to monitor their daily routines and provide timely care when their behavioral movements are identified as abnormal.

In terms of daily health care, dieticians can use the data on human eating habits and daily energy consumption identified by the HAR system to develop diet plans and health recommendations that are appropriate for each individual's situation [17].

3.2 Sports

Physical activity is an activity that has been practiced by humans since birth, and having good exercise habits is very beneficial for both physical and mental health. In particular, in modern society, the proportion of obese people among all humans is on the rise due to various reasons such as work intensity and behavioral habits, and obesity is associated with diabetes, heart disease and even some cancers [18], which have a significant impact on health. In this context, the importance of reducing body weight through physical activity is particularly important.

HAR research has been widely applied in the field of sports, for example, the accelerometer sensor is matched to the human waist, in the recognition of human walking, running, cycling and other common sports above can achieve more than 80% accuracy [19]. In addition, wearable sensing devices loaded with piezoelectric energy harvesters can identify the energy consumed by humans in sports with low power consumption [20].

In fact, the development of HAR research in sports owes much to the small size and low power consumption of accelerometer sensors, which when integrated into wearable devices can be worn on multiple parts of the human body such as the wrist, waist and legs, facilitating more complex human activities such as identification, rowing, cycling and jumping, and are able to maintain an average accuracy rate of over 96% [21].

4. HAR Process

The process of the HAR system is divided into the following stages, namely sensor selection, data pre-processing, feature extraction, dimensionality reduction and classification. Figure 3 illustrates the general process of HAR and the individual steps are described separately in this section.

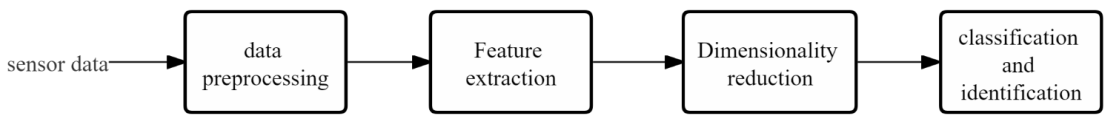


Figure 3 Steps of HAR

4.1 Selecting Sensors

HAR systems generally obtain data through sensors. Different categories of sensors have their own suitable application direction. In the first step of selecting sensors, to fully grasp the advantages and disadvantages of each sensor, appropriate sensor shall be selected. This sub-section mainly lists two categories of sensors commonly used in HAR systems, namely environmental and wearable sensors.

Environmental sensors in HAR research are suitable for fixation in a specific device or environment to sense and recognise human activity in a fixed space. For example, a thermal infrared detector containing an infrared sensor can be fixed in a location where human activity is occurring, allowing the identification of repetitive human activity regardless of the intensity of light or background color [22]. Cameras are a widely used class of sensing device that can capture video containing humans directly, analyze the frames from the video and recognize activity in relation to human silhouette, posture, position and activity time, and are widely used in the medical field [23]. Some studies have sensed human activity through wireless signals emitted from WiFi devices fixed indoors, which, in contrast to camera devices, can penetrate walls and obscured objects, independent of light intensity, and recognize human activity through signals reflected from wireless signals on a moving human body [15].

In the last five years, accelerometer sensors have been one of the most popular sensors in HAR, and many studies have embedded them into wearable devices that are worn on various parts of the human body such as waist, wrist, legs and pockets, sometimes even achieving an average recognition accuracy of 100% in recognizing human activities such as sitting, standing, lying and walking [24]. In addition, due to the popularity of smartphones and the high demand for them by humans, accelerometer sensors integrated into smartphones can sense human activity more comprehensively. For example, a combination of GPS and sensors such as gyroscopes, accelerometers and magnetometers embedded in smartphones can enable navigation functions [25].

Not only should performance issues be fully considered during the sensor selection phase, but other important factors need to be fully taken into account, namely the amount of energy consumed, battery capacity and cost. Many current HAR articles do not take these factors fully into account, but they are very important, especially for wearable sensors, which will create a lot of resistance to human activity recognition if the internal battery of the device does not have enough power to power the components for a sufficient period of time. In order to reduce the amount of energy consumed, on the one hand, short-range wireless networks such as Bluetooth and WiFi can be used as much as possible under the condition that the recognition range is met, avoiding short-range wireless networks such as cellular networks, which are energy-hungry, and minimising the amount of energy consumed by the data during transmission. On the other hand, HAR devices can be loaded with relevant processing elements so that feature extraction and classification can be done within the device, which may reduce the comfort level of the wearable device but contributes more to energy savings by eliminating the need for data transmission. Alternatively, the energy supply in the HAR device can be met by increasing the battery capacity, or by turning off some of the sensors in the device that are not working and then turning them back on to work when they are needed for sampling and transmission, for example, the GPS sensors in the HAR device can be temporarily turned off when a human is recognized to remain standing, sitting or lying down for a long period of time, thus reducing energy consumption [26].

4.2 Data Pre-processing

4.2.1 Denoising

The data obtained from the sensors is often noisy and cannot be analyzed directly; it needs to be pre-processed before the noise can be fed into the classifier. Data pre-processing is a very important stage in the HAR process, and if the noise is not effectively removed, it will greatly affect the average recognition accuracy of HAR.

In some HAR studies, researchers have used low-pass filters to process the acceleration data acquired by accelerometer sensors to separate the high-frequency horizontal acceleration components and the low-frequency vertical acceleration components, after which feature extraction and classification of the high and low frequency data can be performed to obtain an average accuracy of over 91.15% in recognizing human activities such as standing, walking up and down stairs and brushing teeth [27]. Filtering noise acquired from gyroscope and accelerometer sensors can be done efficiently using median filters and Butterworth filters, which eliminate the gravity factor in the vertical direction with a 2hz Butterworth filter [28]. Filtering noise can also be done using a moving average filter to prevent the effect of high frequency signals on the normal signal [29].

In some cases, the sampling rate of the data obtained by filtering WiFi signals is generally low, when the use of traditional low-pass filters does not filter out all the noise, resulting in an unnoticeable distortion in the filtered stream. Some articles have proposed a new scheme for denoising channel state information, which uses principal component analysis to effectively remove high frequency noise [30].

4.2.2 Segmentation

The recognition of human activity often requires it to be continuous in time to facilitate the identification of dynamic human information over time, but in practical research these dynamic, continuous activities are difficult to recognize directly and require the time series data acquired by the sensors to be segmented and then feature extracted. In HAR, segmentation algorithms are often used including sliding window, activity-defined window and event-defined window algorithms.

The sliding window algorithm is one of the most common and important segmentation algorithms used today. It takes the time series acquired by the sensor and segments it into a number of sub-series, each of which is a time window, according to a predetermined length. It is important to note that a very important selection factor for this algorithm is the choice of the length of each time window. In fact, there is no uniform standard for the choice of sliding window size in existing HAR studies, with the shortest being taken as short as 0.08s [31] and the longest being taken as long as 30s or more [32]. Many studies have used sizes chosen from past cases, and these values, which are subject to randomness, are not necessarily optimal choices. The smaller the window, the less information it contains and the faster and more effective the feature extraction. For example, the smaller the window, the less information it contains and the faster and more effective the feature extraction. If the window is too large, it may contain information about another activity, which may affect the feature extraction process later on. For basic activities such as walking and running, a window size of 5.12s is chosen for higher recognition rates when the accelerometer is worn to the waist [33], and a window size of 6s can be chosen when the accelerometer is worn to the wrist, thighs and hips to recognise more complex activities such as shopping and sweeping [34].

When using the sliding window algorithm, the data in the edge portion of each time window is sometimes difficult to extract, and this can be done by overlapping the edge portions of adjacent windows to improve the conversion rate, which can generally be chosen to be 50% [35].

4.3 Feature Extraction

Feature extraction is an important step in the HAR process. After obtaining the pre-processed data, it is usually necessary to process the data to extract the main features and exclude redundant features. The feature vectors extracted after the redundant features have been eliminated are very important and are used directly in the classification algorithm to identify human activity information.

In HAR research, researchers can either manually extract features from raw data or use deep convolutional neural networks in deep learning. Both methods have now been implemented for a wide range of applications with notable success. The advantage of the craft production method is that it can extract both time domain features and frequency domain features together and the computational complexity of the extracted features is not high. The disadvantage is that only data obtained from a few sensors such as accelerometers and gyroscopes can be processed and the person extracting them needs a clear understanding of the data features [36]. In contrast, the advantages of

extracting features through deep neural networks are very obvious, on the one hand, they do not require human involvement and do not require significant human costs, on the other hand, they are less restrictive and can be implemented without domain knowledge and are also very efficient in recognition.

The features extracted in this session are generally time domain features and frequency domain features. In the HAR study more features are extracted in the time domain because it is cheaper to extract time domain features and the time domain can be converted to frequency domain.

4.4 Dimensionality Reduction

Feature sets extracted by hand or deep neural networks often contain many redundant features that are not used for classification, which will lead to over-computation of the classifier and inefficient recognition, i.e. curse of dimensionality. Dimensionality reduction algorithms can effectively solve this problem and can save overhead by removing some of the noise remaining in the dataset. There are a number of methods that can be used to reduce the dimensionality of a feature set. Feature selection is a common method where valuable features are selected directly from the original set of extracted features, thus reducing the dimensionality.

Here it is important to note the difference between feature extraction and feature selection. Feature extraction and feature selection are similar in that they both reduce the dimensionality of a dataset, but they are completely different. Feature extraction focuses on extracting useful features from a feature set and combining features with each other, whereas feature selection is simply a selection that does not change the relationship between features. In feature selection, a suitable subset of features is first selected initially, redundant information is filtered out, and then an efficient evaluation function is built to assess the selection of the feature subset. At the same time, to prevent the evaluation function from being too computationally intensive, a threshold is usually set, and when the function value exceeds this threshold, the function stops searching and evaluating before finally validating the selected features.

Feature selection can be achieved by filtering methods, which generally start by scoring each feature according to the correlation between the original features or between the original features and the target features, and then determining a threshold or filtering by the number of features in the original feature set [37]. Filtering methods are further divided into univariate filtering and multivariate filtering. Univariate filtering is to disregard the interrelationship between the original features and only rank the features by the high or low correlation between the original features and the desired features, and only take the features with high correlation with the target features. This method can generally achieve better computational efficiency and is not prone to situations such as overfitting. Multivariate filtering is ranking only by the high or low correlation between the original features.

Packing and embedding methods are also commonly used in feature selection. In contrast to filtering, which uses machine learning algorithms to select and evaluate feature sets, the packing method eliminates the need for manual operations but is computationally intensive and prone to overfitting. The packing and filtering methods have their own advantages and disadvantages, and the embedding method is a combination of the two, which combines excellent computational efficiency with machine learning algorithms.

4.5 Classification and Identification

For the reduced-dimensional features, the features need to be processed by classification algorithms to obtain activity recognition results. there are generally three types of classification algorithms for HAR, namely traditional machine learning algorithms and deep learning algorithms.

5.5.1 Traditional Machine Learning Algorithms

Traditional machine learning algorithms generally include supervised learning and unsupervised learning algorithms, both of which have a wide range of applications in HAR. A supervised learning

algorithm is a learning task that learns from a pre-determined and labelled training dataset and generates a function, and is often used to solve classification and regression problems.

SVM is a common supervised learning algorithm that is generally used for classification problems and is one of the most effective algorithms available for classification. In implementing classification problems using SVM, one separates the different features inside the variable space by means of a hyperplane. Ahmed et al. [38] used a gyroscope and accelerometer sensor worn around the waist of a volunteer to sense human activity and obtained an accuracy of 96.7% with an SVM classifier.

Artificial neural networks (ANN) are a framework that mimics the neuronal networks of the human brain and are widely used in computer vision, medical diagnosis, etc. Mehr et al.[39] used ANN to monitor human activities such as bathing and walking through a smart home system and could obtain an accuracy rate of 92.81%. Khan et al.[40] used accelerometer sensors to sense human activities such as lying, standing, walking, running and other human activities and could obtain over 99% recognition rate through ANN.

K-Nearest Neighbors (KNN) is a commonly used supervised learning algorithm in which each sample in the feature space is determined by the k samples adjacent to it. fang et al. [41] obtained an average accuracy of 95.3% by KNN and 88.9% by SVM classifier by monitoring the movement of volunteers getting on and off the bus with accelerometers and gyroscopes integrated into the interior of the smartphone average accuracy by KNN and 88.9% by SVM classifier.

Unsupervised learning is a traditional machine learning method that does not require a priori knowledge. K-mean clustering is a commonly used unsupervised learning algorithm that classifies data by determining k as the number of groups in advance. gaglio et al.[42] used an RGB-D camera to sense human joint activity, processed the data using k-mean clustering and obtained 77.3% accuracy. liu et al.[43] proposed a new method that, even without determining the number of activities in advance, can activity perception and recognition within a smartphone, where known human activities can be identified through hierarchical clustering and obtain over 90% accuracy. aljarrah et all.[44] obtained 97.64% recognition efficiency using a PCA-BLSTM model in recognising human activities with wearable sensors. sukor et al.[45] used accelerometer sensors to recognise human activities such as sitting, lying down, walking and walking up and down stairs, the classifier achieved a recognition rate of 96.11% for the frequency domain features after dimensionality reduction of PCA.

5.5.2 Deep Learning

Many traditional machine learning algorithms require manual extraction of features, where deep learning is particularly advantageous, as it can learn deep features and classify them on its own. CNN, RNN, GAN, RL and other algorithms are often used in HAR research.

CNN is a common deep learning algorithm often used for face recognition and image classification. mohmed et al. [46] used environmental sensors such as passive infrared and door sensors to sense human activity within a home environment and obtained an activity recognition rate of 99.5% by a deep convolutional neural network algorithm. ronao et al. [1] proposed a deep convolutional neural network that for activity recognition rate of over 94.79% can be achieved by sensors sensed inside smartphones.

RNNs can be used to process sequential data and are commonly used to implement speech recognition, machine translation, etc. Park et al.[47] analyzed dynamic human joint activity and achieved up to 99.55% recognition accuracy with the RNN algorithm. Moshiri et al.[48] identified indoor human activity through WiFi signals and improved classification accuracy by up to 3.4% compared to the PCA algorithm when using Vishwakarma et al. [49] achieved a maximum of 8% improvement in classification performance through the use of the GAN algorithm when monitoring human activity using WiFi radar.

5. Summary

In this paper, we surveyed the relevant background, research directions and specific applications about HAR in healthcare, sports and other fields in the last decade, described the sensors used in

HAR, and explained the steps of HAR, including sensor selection, data pre-processing, feature extraction, dimensionality reduction, classification and recognition.

With the rapid development of sensing technology, in our survey we particularly noted the advantages of accelerometer sensors in sensing human activity, which not only have a very low cost, but also a small size and high portability, which has already greatly increased the flexibility and range of environments in which HAR can be applied. In future HAR research, wearable sensors, mainly accelerometer sensors, will have a high chance of dominating the mainstream, with applications in healthcare and sports, and are expected to be an important research direction in response to the global trend of an aging population. In fact, HAR can effectively reduce the cost of medical care, and can effectively do care and life assistance for young children and adults, especially the elderly, and will be more deeply applied to the actual life of human beings in the future society, and the future HAR system will also tend to be intelligent, portable and low-cost.

Also, HAR systems based on environmental sensors are not going to be obsolete, as fixed places such as wards and housing still exist. The fixation of sensors in the human environment allows for efficient monitoring and identification of human activity and facilitates human life. It is important to note that the accompanying private data needs to be reasonably protected against privacy breaches and the number of problems they cause. The future potential of WiFi-based HAR systems is considerable, as they do not depend on the intensity of light, they can penetrate through obstructions, and they can effectively address privacy issues.

References

[1] Ronao C A, Cho S B. Human activity recognition with smartphone sensors using deep learning neural networks[J]. Expert systems with applications, 2016, 59: 235-244.

[2] Kim E, Helal S, Cook D. Human activity recognition and pattern discovery[J]. IEEE pervasive computing, 2009, 9(1): 48-53.

[3] Wang A, Chen G, Yang J, et al. A comparative study on human activity recognition using inertial sensors in a smartphone[J]. IEEE Sensors Journal, 2016, 16(11): 4566-4578.

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

[5] Qiu J, Wu Q, Ding G, et al. A survey of machine learning for big data processing[J]. EURASIP Journal on Advances in Signal Processing, 2016, 2016(1): 1-16.

[6] Hastie T, Tibshirani R, Friedman J. Overview of supervised learning[M]//The elements of statistical learning. Springer, New York, NY, 2009: 9-41.

[7] Hastie T, Tibshirani R, Friedman J. Unsupervised learning[M]//The elements of statistical learning. Springer, New York, NY, 2009: 485-585.

[8] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. nature, 2015, 521(7553): 436-444.

[9] Poitras I, Dupuis F, Bielmann M, et al. Validity and reliability of wearable sensors for joint angle estimation: A systematic review[J]. Sensors, 2019, 19(7): 1555.

[10] Yeong D J, Velasco-Hernandez G, Barry J, et al. Sensor and sensor fusion technology in autonomous vehicles: A review[J]. Sensors, 2021, 21(6): 2140.

[11] Li G, Liu S, Wang L, et al. Skin-inspired quadruple tactile sensors integrated on a robot hand enable object recognition[J]. Science Robotics, 2020, 5(49): eabc8134.

[12] Barna A, Masum A K M, Hossain M E, et al. A study on human activity recognition using gyroscope, accelerometer, temperature and humidity data[C]//2019 international conference on electrical, computer and communication engineering (ecce). IEEE, 2019: 1-6.

[13] Kanasi E, Ayilavarapu S, Jones J. The aging population: demographics and the biology of aging[J]. Periodontology 2000, 2016, 72(1): 13-18.

[14] Jiang S, Cao Y, Iyengar S, et al. CareNet: an integrated wireless sensor networking environment for remote healthcare[C]//BODYNETS. 2008: 9.

[15] Khan U M, Kabir Z, Hassan S A, et al. A deep learning framework using passive WiFi sensing for respiration monitoring[C]//GLOBECOM 2017-2017 IEEE Global Communications Conference. IEEE, 2017: 1-6.

[16] Győrbíró N, Fábián Á,Ományi G. An activity recognition system for mobile phones[J]. Mobile Networks and Applications, 2009, 14(1): 82-91.

[17] Arif M, Bilal M, Kattan A, et al. Better physical activity classification using smartphone acceleration sensor[J]. Journal of medical systems, 2014, 38(9): 1-10.

[18] Kopelman P G. Obesity as a medical problem[J]. Nature, 2000, 404(6778): 635-643.

[19] Long X, Yin B, Aarts R M. Single-accelerometer-based daily physical activity classification[C]//2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2009: 6107-6110.

[20] Khalifa S, Lan G, Hassan M, et al. Harke: Human activity recognition from kinetic energy harvesting data in wearable devices[J]. IEEE Transactions on Mobile Computing, 2017, 17(6): 1353-1368.

[21] Khan A M, Lee Y K, Lee S Y, et al. Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis[C]//2010 5th international conference on future information technology. IEEE, 2010: 1-6.

[22] Han J, Bhanu B. Human activity recognition in thermal infrared imagery[C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops. IEEE, 2005: 17-17.

[23] Kim K, Jalal A, Mahmood M. Vision-based human activity recognition system using depth silhouettes: A smart home system for monitoring the residents[J]. Journal of Electrical Engineering & Technology, 2019, 14(6): 2567-2573.

[24] Yeoh W S, Pek I, Yong Y H, et al. Ambulatory monitoring of human posture and walking speed using wearable accelerometer sensors[C]//2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2008: 5184-5187.

[25] Al-Mazloum A, Omer E, Abdullah M F A. GPS and SMS-based child tracking system using smart phone[J]. International Journal of Electronics and Communication Engineering, 2014, 7(2): 238-241.

[26] Reddy S, Mun M, Burke J, et al. Using mobile phones to determine transportation modes[J]. ACM Transactions on Sensor Networks (TOSN), 2010, 6(2): 1-27.

[27] Bayat A, Pomplun M, Tran D A. A study on human activity recognition using accelerometer data from smartphones[J]. Procedia Computer Science, 2014, 34: 450-457.

[28] Bulbul E, Cetin A, Dogru I A. Human activity recognition using smartphones[C]//2018 2nd international symposium on multidisciplinary studies and innovative technologies (ismsit). IEEE, 2018: 1-6.

[29] Adaskevicius R. Method for recognition of the physical activity of human being using a wearable accelerometer[J]. Elektronika ir Elektrotechnika, 2014, 20(5): 127-131.

[30] Wang W, Liu A X, Shahzad M, et al. Understanding and modeling of wifi signal based human activity recognition[C]//Proceedings of the 21st annual international conference on mobile computing and networking. 2015: 65-76.

[31] Berchtold M, Budde M, Schmidtke H R, et al. An extensible modular recognition concept that makes activity recognition practical[C]//Annual conference on artificial intelligence. Springer, Berlin, Heidelberg, 2010: 400-409.

[32] Tapia E M, Intille S S, Haskell W, et al. Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor[C]//2007 11th IEEE international symposium on wearable computers. IEEE, 2007: 37-40.

[33] Ravi N, Dandekar N, Mysore P, et al. Activity recognition from accelerometer data[C]//Aaai. 2005, 5(2005): 1541-1546.

[34] Huỳnh T, Blanke U, Schiele B. Scalable recognition of daily activities with wearable sensors[C]//International Symposium on Location-and Context-Awareness. Springer, Berlin, Heidelberg, 2007: 50-67.

[35] Lara O D, Pérez A J, Labrador M A, et al. Centinela: A human activity recognition system based on acceleration and vital sign data[J]. Pervasive and mobile computing, 2012, 8(5): 717-729.

[36] Sani S, Massie S, Wiratunga N, et al. Learning deep and shallow features for human activity recognition[C]//International conference on knowledge science, engineering and management. Springer, Cham, 2017: 469-482.

[37] Gheid Z, Challal Y. Novel efficient and privacy-preserving protocols for sensor-based human activity recognition[C]//2016 Intl IEEE Conferences on 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). IEEE, 2016: 301-308.

[38] Ahmed N, Rafiq J I, Islam M R. Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model[J]. Sensors, 2020, 20(1): 317.

[39] Mehr H D, Polat H, Cetin A. Resident activity recognition in smart homes by using artificial neural networks[C]//2016 4th international istanbul smart grid congress and fair (ICSG). IEEE, 2016: 1-5.

[40] Khan A M, Lee Y K, Kim T S. Accelerometer signal-based human activity recognition using augmented autoregressive model coefficients and artificial neural nets[C]//2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2008: 5172-5175.

[41] Fang L, Yishui S, Wei C. Up and down buses activity recognition using smartphone accelerometer[C]//2016 IEEE Information Technology, Networking, Electronic and Automation Control Conference. IEEE, 2016: 761-765.

[42] Gaglio S, Re G L, Morana M. Human activity recognition process using 3-D posture data[J]. IEEE Transactions on Human-Machi

[43] Liu A A, Su Y T, Nie W Z, et al. Hierarchical clustering multi-task learning for joint human action grouping and recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 39(1): 102-114.

[44] Aljarrah A A, Ali A H. Human activity recognition using PCA and BiLSTM recurrent neural networks[C]//2019 2nd International Conference on Engineering Technology and its Applications (IICETA). IEEE, 2019: 156-160.

[45] Sukor A S A, Zakaria A, Rahim N A. Activity recognition using accelerometer sensor and machine learning classifiers[C]//2018 IEEE 14th international colloquium on signal processing & its applications (CSPA). IEEE, 2018: 233-238.

[46] Mohmed G, Lotfi A, Pourabdollah A. Employing a deep convolutional neural network for human activity recognition based on binary ambient sensor data[C]//Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments. 2020: 1-7.

[47] Park S U, Park J H, Al-Masni M A, et al. A depth camera-based human activity recognition via deep learning recurrent neural network for health and social care services[J]. Procedia Computer Science, 2016, 100: 78-84.

[48] Moshiri P F, Navidan H, Shahbazian R, et al. Using GAN to enhance the accuracy of indoor human activity recognition[J]. arXiv preprint arXiv:2004.11228, 2020.

[49] Vishwakarma S, Tang C, Li W, et al. Gan based noise generation to aid activity recognition when augmenting measured wifi radar data with simulations[C]//2021 IEEE International Conference on Communications Workshops (ICC Workshops). IEEE, 2021: 1-6.