Journal of Computer Science

Review

Human Activity Prediction Studies Using Wearable Sensors and Machine Learning

Divya Sharma and Usha Chauhan

Department of Electrical, Electronics and Communication Engineering, Galgotias University, Greater Noida, India

Article history
Received: 02-10-2023
Revised: 12-01-2024
Accepted: 17-01-2024

Corresponding Author:
Divya Sharma
Department of Electrical, Electronics and Communication Engineering, Galgotias University, Greater Noida, India
Email: divyasharma@galgotiasuniversity.edu.in

Abstract: Nowadays, Human Activity Recognition (HAR) systems have become so demanding that they find applications in the field of assisted living systems, elderly healthcare systems, smart homes, healthcare monitoring applications, surveillance systems, etc. Due to its increasing demand in computer science; machine learning and deep learning have brought a paradigm shift in the area of sensor-based recognition systems. Providing accurate information about individuals is essential in pervasive computing, but human activity detection is challenging due to the complexity and speed of activities, dynamic recording requirements, and diverse application areas. This survey aims to identify the best wearable device and most optimal machine learning algorithms for HAR in terms of classification accuracy, as well as analyze which algorithms are suitable for specific application areas. The recent advances in HAR systems through machine learning and deep learning techniques have been discussed and from this analysis, it has been observed that CNN and RNN-LSTM techniques have achieved maximum classification efficiency of 92-95.78% on the ADL dataset.

Keywords: Activities of Daily Living (ADL), Machine Learning, Wearable Technology

Introduction

Human Activity Recognition (HAR) is a time series classification task that deals with the development of systems that help in recognizing the actions or behavior of a person from a series of observations mapped against a predefined set of tasks. In the last few years outbreak of coronavirus has affected all the countries of the world. It has been observed that both mental and physical health are being affected in the COVID phase, with almost 16-18% of patients showing symptoms of anxiety and alexithymia. Studies show that people in the age group 18-28 years and with poor quality of sleep are at higher risk for physical and mental health issues. So, proper monitoring of day today's activities is required. Abbas and Jeannès (2021) has discussed to avoid this type of critical situation; one should take care of their own health and one's health by monitoring the biological signals by means of wearable technology. HAR is also crucial for automating surveillance systems to monitor ambient conditions and detect suspicious activity. Anguita *et al*. (2012) has introduced human activities are typically recognized from a sequence of movements captured by vision-based sensors or non-vision sensors. This survey highlights the various active learning techniques for sensor-based activities proposed in the last decade, challenges faced in data acquisition, classification accuracy attained in datasets in terms of fall detection and fall classification, and rigorously provides the comparison of various successfully applied algorithms. Various systems have been reviewed with respect to flexibility, recognition performance, obtrusiveness, and energy utilization. Finally, the Existing survey summarizes seven aspects: Design issues in a HAR system, sensor modalities, battery usage, sensors used in wearable technology, and classification accuracy achieved based on the algorithm used for the problem statements of the last decade. Hence, some of the open research problems are addressed in this research paper which has high relevance in future research.

This study is organized as follows: In the next section, a detailed description of the HAR process is given. After that literature survey, different types of activities recognized by the HAR system, sensor modalities, types of sensors used, market share of wearables, Battery usage by wearables, and in the last section number of papers published on the basis of machine learning and deep learning techniques used in the last decade has been discussed. This survey paper gives a detailed description of the role and importance of the HAR system in real-life applications.

HAR is the task of identifying the actions of a person based on sensor data and contextual information. HAR systems can be approachable in two types-one is through wearable sensors and the second is through external sensors. Meanwhile, the interpretation of these activities is completely dependent upon the spontaneous intercommunication of the user and the external wearable sensors. Banos *et al*. (2014) provided a detailed survey of the research papers published over the past decade shows that the reduction in efficiency of human activity recognition systems is attributed to the following major facts: (i) Number of observations in each second is huge hence data cleaning and processing algorithms lack to process the data with high accuracy rate (ii) Temporal resolution of the observations (iii) Lack of clear vision to relate sensor data to physical conditions (iv) Noise picked up by sensors while recording the physiological activities. Current smartphones and wearable accessories have inbuilt embedded sensors that provide a significant amount of data to measure motion characteristics during these physical activities. Ayu *et al*. (2012) has given an immensely popular example of external sensing is a smart home system which consists of various sensors that can record, analyze, and transduce all kinds of daily living activities. However, these sensors are positioned at various locations, so it will be easy to get data from the focused target. Nevertheless, due to shadows and occlusion re-coding of data is not possible due to the sensor having limited range.

Figure 1 the structure of the HAR system has been shown, where cameras are part of external sensors. These computer vision systems have been used extensively in the security and interactive applications areas. Since image and video processing have some drawbacks (1) Secrecy sometimes people are not willing to be monitored and recorded through cameras. (2) Pervasiveness during observation, it is very difficult to focus or take images of the position of body parts, the person must be monitored under the required dimensions defined by the location of cameras. (3) Intricacy however image processing activities are based on certain computational parameters which makes it a bit expensive. Hence, the above-named constraints inspire the usage of Human activity recognition systems that are sensor-based and wearable. Several types of sensors are accelerometers (used for detecting the user's movement), gyroscope (gives the angular motion) temperature and humidity sensors, heartbeat sensors, etc. The purpose of HAR, based on the investigation of different types of household activities is (a) The principle of learning models that grant the segregation of the normal and abnormal nature or conduct of individuals, (b) To facilitate the needed tools for the custodian and the medical unit to analyze the activities performed by them and develop some good precautionary measures. Human activity recognition systems follow a systematic process in Fig. 2.

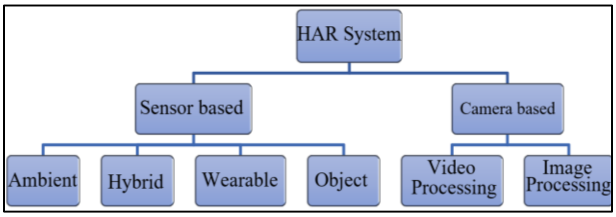


Fig. 1: Structure of HAR system

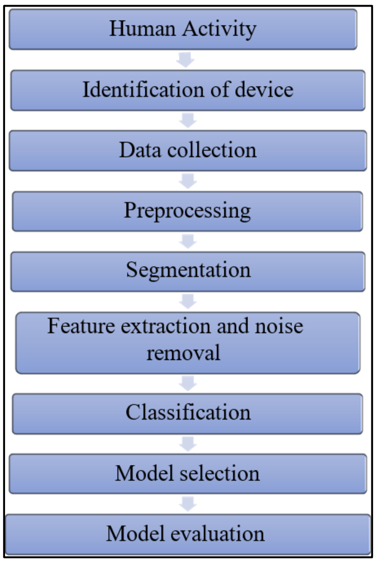


Fig. 2: Standard workflow for implementing HAR-based applications

The information collected from different types of sensors present in the wearable device attached to the body is stored in the form of datasets. The selection of features is a very important step in the process of the HAR system. It will always take place before classification. Dimensions of the feature set must be considered very carefully as it may affect the recognition accuracy. With the selection of a high-dimensional element, it will provide only redundant and inappropriate information. Secondly, it is very difficult to train a high dimensional feature i.e., why it is recommended to use a reduced dimension feature set before performing the classification process.

The first step in developing a HAR application is to check the nature of the sensor and the device used for collecting the information. Secondly, the collection of data is the other step in analyzing such activities. This collected data is then sent for the training process in which it will train the multiple Machine Learning (ML) algorithms which will conclude the human behavior analysis. For example, this may help to send early warning messages to guardians about the illness of their dear ones. Said that due to its ineffectiveness in various applications, many ML algorithms such as decision trees, KNN, support vector machines, etc. are used to improve the system's accuracy and robustness.

According to Cruciani *et al*. (2018), the functioning of the HAR system is shown in Fig. 3. In this system, the user will perform various ADLs and the recording of data like location, heart rate, calories burnt, body temperature, etc., will be recorded by a specific wearable device that can be a smart band or smart clothes then this recorded data will be monitored through a mobile phone application and the technology used for the transmission of data from a wearable device to mobile phone application can be Bluetooth or Wi-Fi. In this way, the whole system will work.

Recent Studies in HAR Systems

Ugulino *et al*. (2012) have explained a complete experimental detail using an ML-based classifier in which a public domain dataset has been recorded from 4 different subjects performing 5 different activities using wearable accelerometers on different body parts and the classification accuracy achieved is 92.43%. Anguita *et al*. (2012) proposed a smartphone-based system using an SVM algorithm that operates with the incidence of postural transitions and has attained a classification accuracy of 91.40%. Khan *et al*. (2021) have implemented a smartphone-based scheme for extracting time-domain quantities and a nonlinear-based approach has been incorporated for the recognition of ADL with 92.70% classification accuracy in both offline and online modes. Fareed (2015) presented the evaluation of different classification algorithms for recognizing eight physical activities using general machine learning algorithms and achieved a classification accuracy of 88.74%. González *et al*. (2015) proposed a novel approach for HAR for monitoring the health issues related to early stroke diagnosis with the help of the most used GFFSM method. Zhu *et al*. (2018) have recommended a unique approach for smartphone-based HAR by using LLC as a feature selector for improving the overall performance of HAR systems and achieving a classification accuracy of 89.86%. González *et al*. (2015) have proposed a unique method for a gymnastic dataset that discriminates normal resting from stroke-related paralysis by using the extension of the GFFSM method and a new (FS) algorithm with PCA. Capela *et al*. (2015) have discussed about the data collected from populations of different types of patients using smartphone accelerometers and gyroscope sensors. The evaluation of the feature subset has been done using three generic classifiers (SVM, decision tree, and correlation-based feature selection).

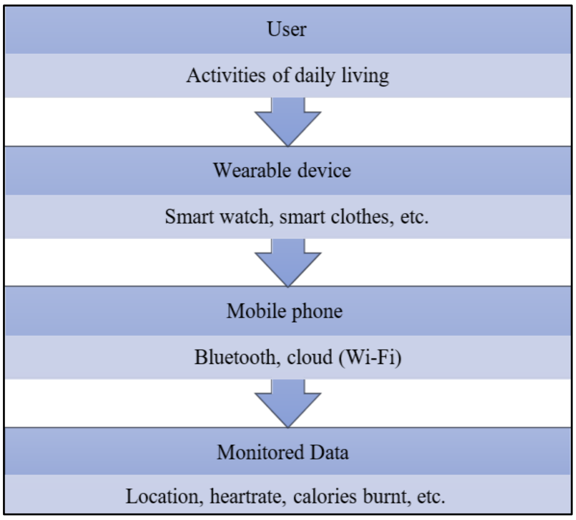


Fig. 3: Basic functional block diagram of the HAR system

Anguita *et al*. (2012) have proposed three machine learning algorithms (HMM, SVM, SVM-HMM) based on their application on elderly people's HAR dataset. Recorded data will serve as an input to a two-way activity awareness system to attain an accuracy of 93.60%. Guo *et al.* (2020) have implemented a novel approach for MARCEL. In MARCEL, the neural network technique has been used as a base classifier for constructing a specific approach, and a combination of multiple classifiers is used as the error function of this technique. Extensive Analysis of the experiments shows that MARCEL yields a varying HAR performance and maximum accuracy. Lee and Kim (2016) have discussed a novel technique for an energy-efficient system by controlling the duration of activity recognition. The investigation has been performed by using the Variable Activity Recognition Duration (VARD) strategy. This technique has been developed on an Android platform and its performance has been analyzed in terms of accuracy and energy efficiency. Results show energy consumption of about 44.23-78.85% without sacrificing the classification accuracy. Ronao and Cho (2016). has proposed a deep convenient technique for performing different human activities effectively and efficiently using smartphone sensors. Experimental results show the relevant and complex features with three additional layers with a classification accuracy of 94.79%. Castro *et al.* (2017) have presented an IoT-based novel HAR system using Machine learning algorithms for the determination of four pre-established activities (lying, sitting, walking, and jogging).

This system provides feedback with remotely programmable alarms, during and after the activity is performed with a classification accuracy of 93.83%. Janidarmian *et al.* (2017) have developed a complete dataset of elderly people with multiple heterogeneities from accelerometer sensors. An extensive analysis has been performed on feature extraction and classification techniques using 293 classifiers. Secondly, PCA has been used for updating the feature vector while keeping the data safe.

Chen *et al*. (2018) have early-stage PD patient data for classifying the sensor signal into an activity, DL based HAR model has been developed for an early-stage PD patient. Their activity profile analysis shows significant differences (walking, standing, and sitting) between PD patients and healthy individuals with an accuracy of 79.76%. Chen *et al*. (2018) have proposed inertial sensors and barometers. This method recognizes 8 classes followed by a multi-layer strategy. The classifiers-based approach has been adopted for different classifications with an accuracy of 77.65%. Bhat *et al*. (2020) have introduced a new framework that will execute both training and assumptions in online mode. The proposed work starts with developed features using FFT and DWT in accordance with accelerometer data. With the help of these features, a neural network classifier has been designed to train in online mode using the PGA with an accuracy of 94.70%. Jain and Kanhangad (2017) have used a descriptor-based approach for the differentiation of different activities using smartphone sensors (accelerometer and gyroscope). In this study, two descriptors, namely, histogram-based Fourier descriptors (multiclass SVM and KNN classifiers), have been used for the extraction of feature sets from these samples and the obtained classification accuracy is 93.83%. Wang *et al.* (2019a-b) have built a deep learning-based automatic feature selection technique and hand-crafted feature selection technique that focuses on how data can be achieved from different types of sensor modalities. Bhat *et al.* (2020) created a wearable dataset by recording seven different activities of 22 subjects. The unique feature of this dataset is the accomplishment of data from wearable and inertial sensors with a classification efficiency of 95%. Agarwal and Alam (2020) have presented a Lightweight DL model for HAR on edge devices. By using the RNN-LSTM Technique. The proposed model is evaluated based on accuracy, Precision, etc., and the attained accuracy is 95.78%. Mekruksavanich *et al*. (2020) created a wearable dataset by recording twelve different ADLs of 43 subjects. The unique feature of this dataset is accomplished by using CNN-LSTM from wearable and inertial sensors have designed a wearable wireless node using a new algorithm IPL-JPDA and it shows the best performance with an improvement of 2.45% among the other five models.

Different Activities Recognized by HAR System

HAR systems recognize several types of activities that can further be classified based on the types of areas. These activities can be distinguished on the basis of four areas i.e., Daily living, Freightage, mobile usage, and gymming. Table 1, briefly summarizes the different activities recognized by HAR systems. Activities of daily living will include most of the natural routine work activities such as walking, standing, sitting, running, lying, climbing stairs, and defending stairs. Riding a bus, cycling, and driving will come in the category of freightage. Calling, messaging, and using any application will come in mobile usage and lastly gymming activities like weightlifting, push-ups, and treadmill.

Currently, the availability of a large number of datasets in HAR makes classification a difficult task for a researcher. HAR systems are well known for their ability to learn the activities of daily living. However, as discussed in the practical applications of HAR systems are countless: Fall detection, energy expenditure estimation, gait anomaly detection, stress detection, behavior monitoring, and rehabilitation. Therefore, the reliability and sustainability of these systems can be attained in terms of their ability to presume the different activities.

Sensor Modalities for HAR system

Despite of many HAR approaches, HAR techniques can be categorized based on the type of sensor modalities. These modalities can be classified based on four aspects i.e., wearable sensors, object sensors, ambient sensors, and hybrid sensors. Table 3 briefly summarizes the different types of sensor modalities.

Wearable Sensors

Wearable technology is the most popularly used method in HAR systems. These sensors (smartphone, smartwatch, smart band's, smart clothes, accelerometer, gyroscope, magnetometer, etc.,) are worn by an individual to record the body movements. The readings measured such as acceleration and angular velocity are recorded to infer the activities of human body movements. Wearables with sensors monitor how the body is moving giving a better understanding. The reason for measuring the activity can be related to a person's desire to gather information in an area where it can work. Advances in sensors change one's lifestyle i.e., a person can gain more control of himself.

Table 1: Different activities recognized by HAR systems

Activity types	Area
Standing, walking, sitting, lying, running, climbing stairs, descending stairs	Daily living
Driving, cycling, riding	Freightage
Calling, text messaging, using any application	Mobile usage
Weightlifting, spinning, push-ups, treadmill	Gymming

Table 2: Sensor modalities for the HAR system

Methods	Characteristics	Examples
Wearable sensors	Wearable by an individual at any part of the body represent the different body movements	Smartwatch, smart bands, smartphone smart outfits, accelerometer, gyroscope, magnetometer, goggles
Object sensors	Connected to objects for recording objects movements	RFID tags, accelerometer on items, gyroscope on belts
Ambient sensors	Enforced in the environment for gathering information from user and climatic sensors interaction	Sound, Wi-Fi, radar, temperature sensor, Bluetooth door
Hybrid sensors	Intersecting the boundaries of sensors with multiple sensors	a combination of different types of sensors often used extensively in smart environments

Table 3: Different types of design issues

Design issues	Types
Selection of attributes or sensor	Environmental, acceleration, location, physiological signals
Intrusiveness	Cell phone, laptop
Data collection protocol	UDP/IP, TCP/IP, cellular network, Wi-Fi
Recognition performance	Location and remote server
Energy consumption	
Processing	
Flexibility	

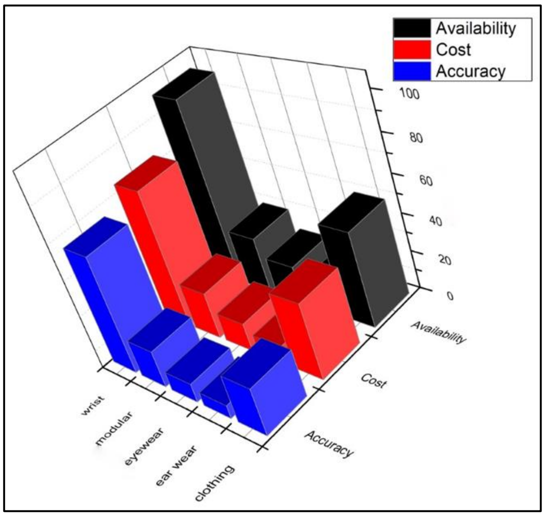


Fig. 4: Classification of wearables in terms of cost-effectiveness, accuracy, and availability

Figure 4 shows that wristwear will be the best choice in terms of accuracy, cost, and availability. As a maximum number of users prefer smart wearables, manufacturers will work to improve in this segment but depending upon the positive consumer response, another body part can be used for different perspectives and device placements.

Object Sensors

For inferring human body movements, sensors are employed on certain objects. These are different from wearable sensors. For example, an accelerometer connected to a glass of water can be used to detect the amount of water in the glass. Similarly, RFID tags are extensively used as object sensors in various applications of healthcare monitoring systems. Due to their deployment difficulties, these sensors are in less use as compared to wearable sensors.

Ambient Sensors

Ambient sensors can also be called climate sensors as they are useful in recording the changes in environmental conditions. For gathering the interaction between an individual and the surrounding environment, the most used sensors are radar, sound, pressure, and temperature sensors. As they are easily influenced by the environment, only some specific applications can make use of these sensors.

Hybrid Sensors

Hybrid means the combination of different varieties of sensors. To improve the accuracy of data collection in a HAR system, some sensors are combined with other sensors to retrieve more data. For example, an ambient sensor embedded with object sensors records the object's movements and the state of the environment. Smart homes consist of multiple sensors, hence multiple data can be collected through wearable, object, and ambient sensors. Hence, Hybrid sensors provide rich information as compared to other sensors.

Sensors Used in the HAR System

Figure 5 shows the interconnection of seven different types of sensors with microcontrollers. These sensors are generally used in wearables for monitoring the activities of daily living. The functioning of these sensors has been discussed briefly in the next section.

Microcontroller

The critical part in permitting wearable innovation to work is a microcontroller. This is ordinarily seen as a smaller than normal PC (framework on a chip). The most fundamental part expected for the working of wearable innovation is a microcontroller. This is a framework on a chip. It empowers the utilization of the Internet of Things (IoT), and it diminishes the need for different electronic parts that make the framework massive. It has its application in wearable innovation in view of its effortlessness, straightforwardness to program, cost, size, similarity with different sensors, and capacity to control complex results, the microchip takes the information and cycles the information as indicated by the program.

DOI: 10.3844/jcssp.2024.431.441

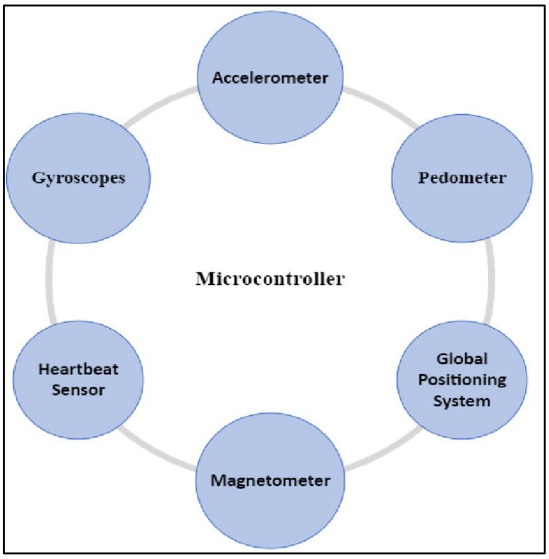


Fig. 5: Interconnectivity of multiple sensors in the HAR system

Accelerometer

Another commonly found sensor in wearables is accelerometers, their ability to allow the data to be reprogrammed for different users is a merit of using this sensor. This sensor can measure the top speed and acceleration when a user runs, this can also monitor the sleep patterns. These applications show that the industries working in the field of sports and medical fitness can benefit from these wearables.

Gyroscopes

Gyroscopes are also used as sensors in wearables. To measure the angular and rotational accelerations, gyroscopes are used. Gyroscopes detect angular velocity on their disk. The vibration gyroscope measures angular velocity because of the Coriolis force. The motion induces a potential difference that is converted into an electrical signal. Gyroscopes measure orientation as well as projection, this can be beneficial in data collection.

Magnetometers

Magnetometers combined with accelerometers and gyroscopes form the Inertial Measuring Unit (IMU). Every one of these sensors can have three axes each, contingent upon the sort. It is basically the same as what a compass does and it assists with coordination. While it is ordinarily utilized with the other two sensors, it supplements them by shifting the direction of the developments.

Global Positioning System (GPS)

GPS is commonly used in various devices. It is used to locate a device or a user. The data collected from the user's device is sent to the satellite where the exact location with timestamp is calculated. This process is similar to the transmission and reception of this system. It is used in smart devices to measure important parameters such as precise location and distance. This system has a wide range of applications with sports being most prominent.

Heart Rate Sensors

Heart rate sensors are a revolution in the field of biosensing, these have provided a sustainable and efficient method to measure the heart rate levels of a human, Photoplethysmography is the process that utilizes light energy to measure the blood flow in veins hence linking this to heart beats. This sensor mainly depends upon the use of a photodiode. The change in wavelengths of the light determines the penetrating capacity through the tissues.

Pedometers

Pedometers are used commonly in today’s world mostly in smart watches or smart wearables.

The basic purpose of this is to count the number of steps walked by a user. Electrical pedometers are the most commonly used these days which rely on MEMS for better accuracy but the working principle is the same as that of a mechanical pedometer. The mechanism of a pendulum is incorporated to calculate the number of steps. With every step to and for motion of a pendulum begins and ends.

Pressure Sensors

Pressure sensors work on the principle of strain gauges, on applying force on the sensor some amount of resistance is produced. These sensors can be improved by using graphene-based flexible sensors, which measure the change in graphene conductivity depending on the resistance. Using a flexible sensor is helpful as it can be incorporated into smart clothing and can make a more flexible wearable. Barometric pressure sensors are also abundantly used in smart wearables. They measure atmospheric pressure relative to the environmental conditions. These have a great application when the need to measure the body conditions of a person at different levels is greater.

Design Issues in HAR Systems

Despite many challenges in HAR systems, several types of design issues are discussed in the above-listed Table 3. Issues such as the selection of attributes or sensors will be due to environmental, acceleration, location, and physiological signals. Issues related to cell phones; and laptops will create intrusiveness. Sometimes issues may come due to data collection protocol or due to wireless technologies such as UDP/IP, TCP/IP, cellular network, and Wi-Fi. Location and remote server errors will also create problems in recognition performance, energy consumption, processing, and flexibility.

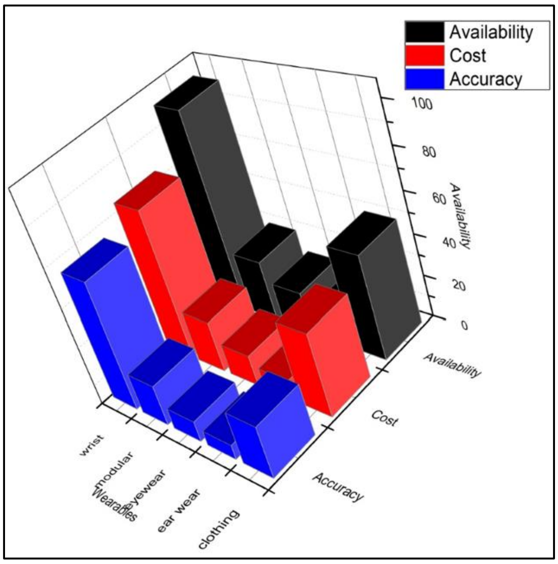


Fig. 6: Percentage-wise global market share of wearables

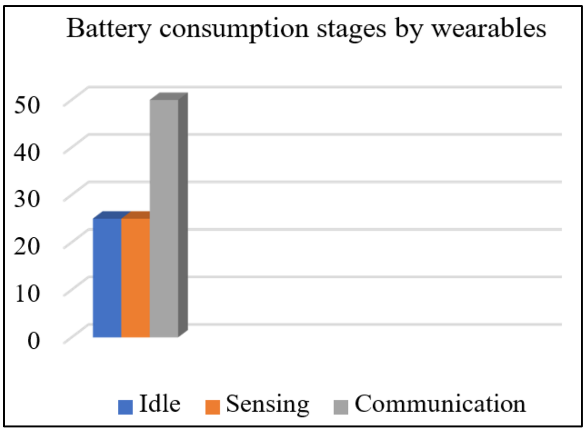


Fig. 7: Percentage of battery consumption states (Idle, sensing, communication) in wearables

Market Share of Wearable Technology

The global market share of these wearable devices has been shown percentage-wise in the Fig. 6. In this graph the data has been represented in terms of sales versus the market share. From the previous research papers, it has been found that the medical industry has having 22% share, the industrial has a 16% share, the defense industry has 27%, consumer usage has the maximum weightage of 72% and others have having 9% share. These Consumers include the young youth, the heart patients, the caregivers, and the elder generation.

Battery Usage by Wearables

Battery plays a very important role in the wearables. The main source of power in these smart devices are batteries be it lithium-ion or solar-powered cells. Further, these batteries should be small in size so that the wearable should not become bulky. The battery consumption is shown in Fig. 7. Representing 0-25% if the device is in an idle state, 25% for sensing state and it should use 50% of the power when in use for communication.

Concept of Machine Learning in HAR Systems

The data gathered from dissimilar sensors deployed in the cloud or from wearable sensors are hoarded in the form of datasets. This dataset is trained by different ML techniques that will predict the behavior of individuals, the purpose of these is to send early warnings to and assess the risk of deteriorating health of the people under observation. To design an ML application, activity recognition consists of two classes, that is training class and the testing class. Training class has a data set obtained from a set of different people performing different daily living activities. This time series classification task is divided into n number of time series window functions for applying feature extraction to filter out the relevant data from the available raw dataset. After that, an ML technique is required to establish a model for the extracted features. Similarly, the testing class contains the data collected from the window functions to extract different features and that feature set is analyzed in the ML model will further generate labels for all the predicted activities. Machine learning is classified into two categories supervised learning techniques.

Unsupervised Learning Techniques

Supervised learning techniques are classified on the basis of regression and classification whereas unsupervised learning techniques are classified on the basis of clustering and association. The few most commonly used algorithms used by researchers in the last five years have been discussed below.

Linear Regression

This technique is used to assess genuine qualities like house cost, number of calls, absolute deals, and so on in the presence of ceaseless variables. Here, one can find out the connection between autonomous and subordinate factors by fitting the best line. This is called the relapse line. This is addressed through a straight-line condition as Eq. 1:

$$Y = mX + c \tag{1}$$

Here, '*y*' and '*x*' represent the dependent and Independent variables respectively, and '*m*' and '*c*' denote the slope and intercept constant.

Types:

- Simple linear regression: It is portrayed by one free factor
- Multiple linear regression: Multiple linear regression is portrayed through different (multiple) free factors.

While observing best best-fit line, one can fit a polynomial or curvilinear relapse

Logistic Regression

This method is used to assess discrete qualities for example binary qualities like 0/1, yes/no, valid/misleading, and true/false. In the presence of certain free variables set. It actually predicts the likelihood of an event by fitting information to logit work. Subsequently, known as logit relapse. Since, it predicts likelihood, therefore, its result lies somewhere in the range of 0 and 1.

Decision Tree

This is the most used supervised learning algorithm. Apparently, it works for dependent variables. In this technique, the data is split into two homogeneous sets and further, the decision will be made based on.

Support Vector Machine (SVM)

SVM is another important classification technique. SVM separately data item is plotted as a point in n-dimensional space with the value of separate feature being assigned as the value of a particular coordinate where n is the number of features.

Naive Bayes

This classification technique is based on Baye's theorem and is easy to implement for large-size datasets. This classifier presumes that the presence of a particular feature in a class is correlated with any of the other available features. Such as some fruit may be a mango if it is yellow and oval in shape. Although these features are dependent upon each other, a naive Bayes classifier would correlate all of these features to wisely contribute to the result that this fruit is a mango. The Eq. 2 is mentioned below is representing the Baye's algorithms as:

$$p(c/x) = p(x/c)p(c)/p(x) \tag{2}$$

where:

- $P(c/x)$ = Posterior probability of class and the given predictor
- $P(c)$ = Prior probability of class
- $P(x/c)$ = Likelihood (probability of predictor given class)
- $P(x)$ = Prior probability of predictor

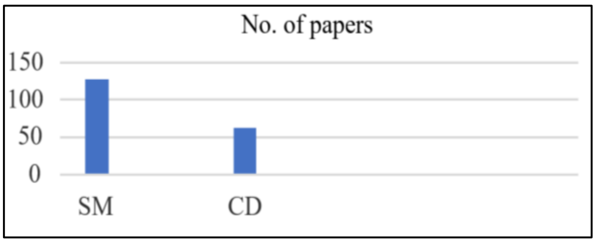
k-Nearest Neighbors (kNN)

This technique is used in classification and regression. kNN works on the principle of collecting and storing all currently available subjects and then classifies a new subject by the majority voting of its k neighbors. It is the simplest algorithm. The subject being allotted to a class is most common amongst its *K* nearest neighbors measured by a distance function. The distance functions for continuous functions are Euclidean and for categorical variables is Hamming distance. A case is assigned to a class of its nearest neighbor when $k = 1$. Sometimes while performing kNN modeling, assuming the value of *K* turns out to be a challenging problem.

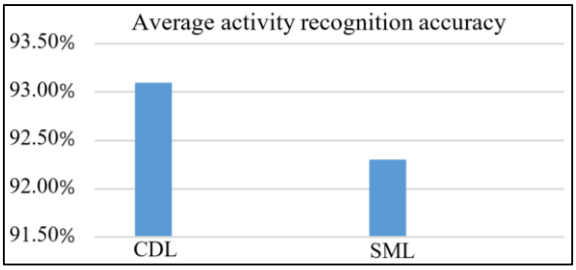
Papers Published on the Basis of Techniques Used in Last Decade

When algorithmic incorporation is considered, there has been a significant rise in Classic Deep Learning methods (CDL), which increases the recognition precision. For a smaller dataset Standard Machine Learning (SML) is better fitted than CDL, lower dimension of input data. The rising interest in HAR can be deduced from the number of publications in the past 10 years as Fig. 8(a) shows that a total of 63 papers were based on CDL models and 127 on SML models and a total of 190 papers were published from 2012-2021. In that period 66 surveys and 24 articles proposed methodologies based on non-ML methods. Fig. 8(b) depicts the mean accuracy among 93 CDL-based papers and 97 SML-based papers (93.1% for CDL-based and 92.3 % SML-based) present nearly similar quality.

Figure 9 last 10-year published HAR paper is reviewed in terms of SML and CDL models, which depicts that the number of SML-based HAR was greater than CDL-based HAR models reason being that all the researchers have taken small-size dataset. The authors have provided both CDL and SML-based methodologies.



(a)



(b)

Fig. 8: (a) Percentage Status of published papers in HAR, based on CDL vs. SML implementations; (b) Their accuracy of recognition of published papers based on CDL vs SML implementations

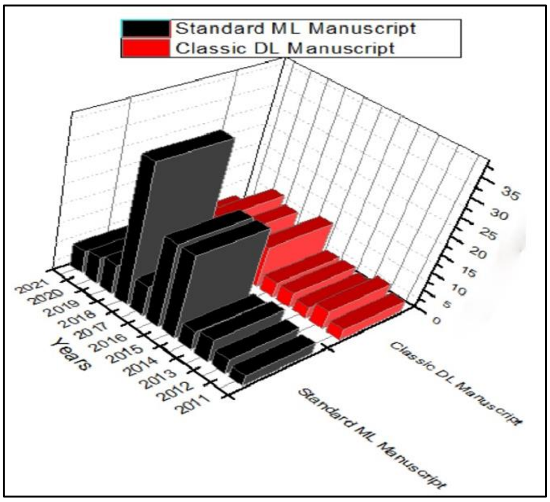


Fig. 9: Percentagewise count of articles published on the basis of SML and CDL techniques

Conclusion

HAR stands as an active focal point within the realm of computer science, spanning its utility across a multitude of domains. HAR data is gathered through the utilization of both visual and non-visual sensors, each bearing relevance and sustainability for specific application areas. This comprehensive survey delves into various machine learning methodologies for the recognition of human activities, encompassing SVM, decision trees, Naive Bayes, KNN, and regression techniques. A detailed analysis of these traditional methods indicates that SVM, decision trees, and KNN exhibit optimal performance across a broad spectrum of application areas. HAR in wearables has enabled us to monitor and improve our health and well-being in many ways. This study integrates the rising use of wearables along with Machine learning and deep learning techniques which utilize sensor data and achieve valuable and distinguished information-based systems. From this survey, it can be concluded that wearable technology is the best choice of HAR modality because after COVID-19, 72% of civilians have started using these gadgets to monitor their day-to-night activities in which wristwear has shown the results with maximum accuracy and availability. The recent advances in HAR systems through machine learning and deep learning techniques have been discussed and from this analysis, it has been observed that CNN and RNN-LSTM techniques have achieved maximum classification efficiency of 92-95.78% on the ADL dataset. Despite the extensive research in HAR, finding the right algorithm and sensor combination for a specific application area remains a challenge.

This survey provides a valuable resource for researchers, offering insights into the current trends and research techniques in the field of HAR. It assists researchers in selecting the most appropriate devices and algorithms for their specific application areas.

Acknowledgment

Authors gratefully acknowledge the support of the department of Electrical, electronics and communication engineering, Galgotias University for their valuable support.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

Divya Sharma: On sensor-based HAR systems.
Usha Chauhan: Reviewed this content.

Ethics

The authors declare that this manuscript has no conflict of interest with any other published source.

Compliance with Ethical Standards

Disclosure of Potential Conflicts of Interest

The authors declare that this manuscript has no conflict of interest with any other published source and has not been published previously (partly or in full). No data have been fabricated or manipulated to support our conclusions.

Research Involving Human Participants and/or Animals

This article does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent

Informed consent was obtained from all individual participants included in the study.

Data Availability Statement

The data that support the findings of this study are available at the reasonable request of the first author.

References

Abbas, M., & Jeannès, R. L. B. (2021). Exploiting local temporal characteristics via multinomial decomposition algorithm for real-time activity recognition. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1-11. https://doi.org/10.1109/TIM.2021.3108235

Anguita, D. Ghio, A. Oneto, L. Parra, X. & Reyes-Ortiz, J. L. (2012). Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. https://doi.org/10.1007/9783-642-35395-6_30

Agarwal, P., & Alam, M. (2020). A lightweight deep learning model for human activity recognition on edge devices. *Procedia Computer Science*, *167*, 2364-2373. https://doi.org/10.1016/j.procs.2020.03.289

Ayu, M. A., Ismail, S. A., Matin, A. F. A., & Mantoro, T. (2012). A comparison study of classifier algorithms for mobile-phone's accelerometer based activity recognition. *Procedia Engineering*, *41*, 224-229. https://doi.org/10.1016/j.proeng.2012.07.166

Banos, O., Garcia, R., Holgado-Terriza, J. A., Damas, M., Pomares, H., Rojas, I., ... & Villalonga, C. (2014). mHealthDroid: A novel framework for agile development of mobile health applications. In *Ambient Assisted Living and Daily Activities: 6th International Work-Conference, IWAAL 2014, Belfast, UK, December 2-5, 2014. Proceedings 6* (pp. 91-98). Springer International Publishing. https://doi.org/10.1007/978-3-319-13105-4_14

Bhat, G., Tran, N., Shill, H., & Ogras, U. Y. (2020). W-HAR: An activity recognition dataset and framework using low-power wearable devices. *Sensors*, *20*(18), 5356. https://doi.org/10.3390/s20185356

Castro, D., Coral, W., Rodriguez, C., Cabra, J., & Colorado, J. (2017). Wearable-based human activity recognition using an IoT approach. *Journal of Sensor and Actuator Networks*, *6*(4), 28. https://doi.org/10.3390/jsan6040028

Capela, N. A., Lemaire, E. D., & Baddour, N. (2015). Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly and stroke patients. *PloS One*, *10*(4), e0124414. https://doi.org/10.1371/journal.pone.0124414

Chen, Z., Jiang, C., & Xie, L. (2018). A novel ensemble ELM for human activity recognition using smartphone sensors. *IEEE Transactions on Industrial Informatics*, *15*(5), 2691-2699. https://doi.org/10.1109/TII.2018.2869843

Cruciani, F., Cleland, I., Nugent, C., McCullagh, P., Synnes, K., & Hallberg, J. (2018). Automatic annotation for human activity recognition in free living using a smartphone. *Sensors*, *18*(7), 2203. https://doi.org/10.3390/s18072203

Fareed, U. (2015, June). Smartphone sensor fusion based activity recognition system for elderly healthcare. In *Proceedings of the 2015 Workshop on Pervasive Wireless Healthcare* (pp. 29-34). https://doi.org/10.1145/2757290.2757297

González, S., Sedano, J., Villar, J. R., Corchado, E., Herrero, Á., & Baruque, B. (2015). Features and models for human activity recognition. *Neurocomputing*, *167*, 52-60. https://doi.org/10.1016/j.neucom.2015.01.082

Guo, J., He, H., He, T., Lausen, L., Li, M., Lin, H., ... & Zhu, Y. (2020). Gluoncv and gluonnlp: Deep learning in computer vision and natural language processing. *The Journal of Machine Learning Research*, *21*(1), 845-851. https://doi.org/10.48550/arXiv.1907.04433

Janidarmian, M., Roshan Fekr, A., Radecka, K., & Zilic, Z. (2017). A comprehensive analysis on wearable acceleration sensors in human activity recognition. *Sensors*, *17*(3), 529. https://doi.org/10.3390/s17030529

Jain, A., & Kanhangad, V. (2017). Human activity classification in smartphones using accelerometer and gyroscope sensors. *IEEE Sensors Journal*, *18*(3), 1169-1177. https://doi.org/10.1109/JSEN.2017.2782492

Khan, P., Kader, M. F., Islam, S. R., Rahman, A. B., Kamal, M. S., Toha, M. U., & Kwak, K. S. (2021). Machine learning and deep learning approaches for brain disease diagnosis: Principles and recent advances. *IEEE Access*, *9*, 37622-37655. https://doi.org/10.1109/ACCESS.2021.3062484

Lee, J., & Kim, J. (2016). Energy-efficient real-time human activity recognition on smart mobile devices. *Mobile Information Systems*, *2016*. https://doi.org/10.1155/2016/2316757

Mekruksavanich, S., Jitpattanakul, A., Youplao, P., & Yupapin, P. (2020). Enhanced hand-oriented activity recognition based on smartwatch sensor data using LSTMS. *Symmetry*, *12*(9), 1570. https://doi.org/10.3390/sym12091570

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. https://doi.org/10.1016/j.eswa.2016.04.032

Ugulino, W., Cardador, D., Vega, K., Velloso, E., Milidiú, R., & Fuks, H. (2012). Wearable computing: Accelerometers' data classification of body postures and movements. In *Advances in Artificial Intelligence-SBIA 2012: 21th Brazilian Symposium on Artificial Intelligence, Curitiba, Brazil, October 20-25, 2012. Proceedings* (pp. 52-61). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-34459-6_6

Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019a). Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, *119*, 3-11. https://doi.org/10.1016/j.patrec.2018.02.010

Wang, Y., Cang, S., & Yu, H. (2019b). A survey on wearable sensor modality centred human activity recognition in health care. *Expert Systems with Applications*, *137*, 167-190. https://doi.org/10.1016/j.eswa.2019.04.057

Zhu, D., Cai, C., Yang, T., & Zhou, X. (2018). A machine learning approach for air quality prediction: Model regularization and optimization. *Big Data and Cognitive Computing*, *2*(1), 5. https://doi.org/10.3390/bdcc2010005