Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 June 2021 doi:10.20944/preprints202102.0349.v2

Human Activity Recognition with Deep Learning: Overview, Challenges & Possibilities

Pranjal Kumar*, *DoCSE, NIT Hamirpur, pranjal@nith.ac.in;*

Abstract

Human Activity Recognition (HAR) has become a vibrant research field over the last decade, especially because of the spread of electronic devices like mobile phones, smart cell phones and video cameras in our daily lives. In addition, the progress of deep learning and other algorithms has made it possible for researchers to use HAR in many fields including sports, health and well being. HAR is, for example, one of the most promising resources for helping older people with the support of their cognitive and physical function through day-to-day activities. This study focuses on the key role machine learning plays in the development of HAR applications. While numerous HAR surveys and review articles have previously been carried out, the main / overall HAR issue was not taken into account, and these studies concentrate only on specific HAR topics. A detailed review paper covering major HAR topics is therefore essential. This study analyses the most up-to-date studies on HAR in recent years and provides a classification of HAR methodology and demonstrates advantages and disadvantages for each group of methods. This paper finally addresses many problems in the current HAR subject and provides recommendations for potential study.

Learning (artificial intelligence); Neural networks; Activity recognition; Multimodal sensors

I. Introduction

Human Activity Recognition (HAR) has been a common subject over the last decade, given its prominence in many fields, including health care, digital gaming, sport, and overall monitoring systems [1]. In addition, the ageing population is now becoming one of the main problems of the nation. Demrozi et al. depict in their study that it has been projected that by 2050, people over 65 years of age would rise from 461 million to 2 billion [2]. This dramatic increase would have important social and health implications. HAR is an important tool for controlling the physical, functional, and cognitive wellbeing of older adults at home [3].

HAR's objective is to consider human activity in regulated and unregulated environments. Despite numerous applications, HAR algorithms still face many challenges, including 1) complexity and variety of day-to-day activities, 2) intra-subject and inter-subject variability for the same operations, 3) trade-off between privacy and performance, 4) embedded and portable system calculation reliability, and 5) data annotation difficulty [4]. Data is normally obtained from two key sources for HAR training and testing, ambient sensors and embedded sensors. Ambient sensors such as temperature sensors or video cameras placed in specific points in the environment can be ambient sensors [5], [6]. Embedded sensing is integrated into or built into clothes and other specialised medical devices, for instance smart phones and smart clocks [7]–[10]. The HAR applications are widespread in the use of cameras, but video data collection presents a wide range of privacy and computational issues [11]. While video cameras generate rich contextual information, the limitations in confidentiality have led many researchers to work with additional environmentally friendly and integrated sensors, including depth images. The standard workflow for developing HAR methodologies is shown in Figure 1.

HAR research has seen an explosion in the use of Deep Learning (DL) methods as an algorithm of implementation, resulting in an improvement in accuracy of recognition [6], [8]. In many HAR applications Classic Machine Learning (CMLs) models can be better adapted due to the limited data set size, the less input data dimension and the availability of expert know-how when formulating the problem [12]. In many HAR applications, DL methods yield highly accurate results with larger activity datasets. The rising interest in HAR can be linked to the increasing use of wearable sensors and appliances in all aspects of everyday life, especially for applications of health and well-being.

II. Existing Surveys

Since HAR is a significant subject of research in recent years, numerous investigations have been published. Among the initial 293 published papers that we reported, 46 were survey papers published since 2015 [2]. Based on data resources

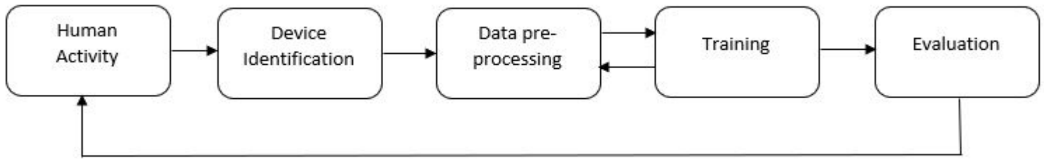

Fig. 1. Workflow for HAR-based applications.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 June 2021 doi:10.20944/preprints202102.0349.v2

and the algorithm of activity recognition, present survey papers can be classified. A), physiological, environmental and inertial devices and b) video recording systems are the most frequently used data sources. With respect to the HAR algorithm, the most commonly used algorithms are CML and DL versions. Our survey paper offers a unique contribution to the literature review by offering a broad view of HAR research development over the last five years. Contrary to current surveys, we don't concentrate solely on the algorithmic information, but we will also identify the data sources in this context (even sensors and devices). We are especially interested in accelerometer sensors, as they demonstrate excellent results in HAR applications and because their use, together with other sensors such as physiological sensors or ambient sensors, is increasingly increasing. The proliferation of sensors is very closely linked to their ability to track the orientation of the human body directly. In addition, it can be affordably incorporated into most wearable devices with accelerometer sensors. The literature was recently reviewed, based on three aspects: sensor modality, DL models and scenarios for use, by Wang. J and colleagues [13] (2019) to include basic details about the work being reviewed. Wang.Y and colleagues [3] (2019) are presenting in the HAR state-of-the-art modalities for sensors, data preprocessing, feature learning, classification, operation, including both the traditional and DL methods, with a focus on the techniques associated with each HAR stage. In addition, HAR, which includes camera-based sensors and systems that combine wearable and ambient sensors, is presented. Sousa Lima et al. (2019) provide a comprehensive, state-of-the-art overview of the current HAR solution for smartphone inertial sensors and Elbasiony and Gomaa [14] (2019) include a thoroughness survey of several HAR systems in portable sensors (Accelerometer, Gyroscope and Magnetometer). D.Triboan et al. [15] reports an approach to fine-grained AR which combines multimodal data from individual objects and addresses the imprecise nature of non-binary sensor measurements. This approach uses fuzzy ontology to model fine-grained behaviour, with imprecise Member States of the sensors, method for classifying action completion with entity and fuzzyDL reasoning. A non-intrusive, heterogeneous ambient and embedded object-based sensing approach is also proposed for the microserver architecture. Darpan Triboan et al. provide semiotic theory based on the ontological model, capturing generic information and residential expectations for the performance of ADLs to help the segmentation process [16].

The comprehensive study of fusion data/sensors and multiple HAR classification systems with focus on mobile and wearable devices is given by Nweke et al. [17] (2019). In the study of 53 papers on physiologic sensors in medical applications including electromyography (EMG), electrocardiogram(ECG), electrooculograms (EOG), and electroencephalograms(EEG), Faust et al. [18] (2018). An overview of ML and data mining techniques used in Activity Recognition (AR), which empathises with fundamental problems and challenges, was provided by Ramasamy Ramamurthy and Roy [19] (2018). Finally, Morales and Akopian [20] (2017) offer an overview of state-of-the-art signals, data collection and preprocessing, on-body positions and orientation calibration, the choice of the right set of functions, the modelling and classification models and how the HAR system can be evaluated. In addition, it includes repetitive movements, postures, falling and inactivity identification. Most of these HAR reviews concentrate on methods of data management and model recognition. As far as we know, no existing survey article provides: (1) a detailed meta-review of the current surveys, (2) a comprehensive overview of different sensors, (3) performance measurement reporting and comparison and (4) reportings on availability and popularity of datasets. [21] presents a knowledge-based approach to the continuous detection of activity based on multi-sensor data in intelligent homes. In [22] an exhausting survey is presented to investigate how different aspects of sensor activity recognition evolve and current status. Models of activities play an important role in the awareness and support of activities in an environmentally supported living environment. A variety of issues, such as cold starting, model reusability and incompleteness, affect existing methods of activity modelling. [23] implement a hybrid ontological approach to Activity Modeling in order to deal with these issues that combines domain knowledge-based modelling and data-driven model learning.

III. Background

HAR algorithms are mainly aimed at acknowledging human activity based on data collected from wearable and environmental sensors [13], [24]. These behaviours are primarily recognised on the basis of CML and DL. In recent times there has been an interest in sensor fusion methods by using a large array of sensors. This section discusses basic ML and DL principles, market evolution and sensor fusion strategies for wearable/environmental sensors.

A. Context of Machine Learning

Machine Learning (ML) is an artificial intelligence branch that has been developed to develop algorithms which can deduce patterns from a training dataset [25]. Two main classes of these algorithms:

- Supervised learning
- Unsupervised Learning

The aim of supervised study is to develop and forecast future unseen data points using a mathematical model based on the relation between the incoming and the output data. The objective is to classify input-data patterns without knowing the result in unsupervised learning [5]. One or more stages of pre-processing, including the extraction, vectorization/segmentation, standardisation or normalisation, and projection, are usually often essential [26].

Some of the most popular CML controlled algorithms are: Naïve Bayes (NB), K-Means Clustering, Support Vector (SVM), Linear Regression, Logistic Regression, Random Forests (RF) (k-NN). By sorting them according to features/data values, DT

 doi:10.20944/preprints202102.0349.v2

classifies data instances. Each node is a categorised function and each branch is a value that the node will assume. Bayes' theorem with strong independence assumptions is founded upon application of NB classification systems. SVMs are based on the idea of a hyperplane on either side that divises two groups of data. To maximise the margin and thus create the maximum distance between a separating hyperplane and instances on both sides, an upper limit on the predicted generalisation error has been proved to be reduced. Finally, K-NN is a CML algorithm that stores and classifies all available cases by similarity (e.g. distances including Euclidean, Manhattan, Minkowski) [25]. In addition, since HAR has unique limitations such as latency reduced, memory restrictions and computational restrictions, these classifiers, with the exception of the SVM, are suitable to low-resource environments due to their low requirements in computer and storage.

The most common algorithms are k-means, hierarchical clusters and mixture models, amongst uncontrolled and especially clustering algorithms. The aim of K-Means clustering is to separate sample groups into K clusters based on the similarity (intra-group) and difference measurements (inter-groups). Each sample is part of the cluster with the closest cluster centre or centroid cluster and serves as a prototype cluster. Hierarchical cluster analysis is a method of cluster analysis aimed at building a hierarchy of clusters in which clusters are combined/split based on measurement of the difference between sets. A mixture model is a probabilistic model that reflects observer subpopulations in the total population [5]. These techniques are especially suitable for processing label-free sets of data, or if a primary outcome [27]–[29] is to quantify similarity/differences between groups.

B. Context of Deep Learning

Many prior studies have implemented methods of machine learning in consideration of human activity [4]. We rely heavily on techniques of abstraction, including transformation of time-frequencies [30], mathematical approaches [31], and symbolic representation [32]. The derived properties are nevertheless carefully developed and heuristic. There were no standardized or systemic methods to derive distinguishable characteristics for human activitys effectively.

In recent years, in many areas of computational vision, natural language processing and voice analysis, deep learning has increased prominence in modeling high-level abstractions of nuanced data [33]. Following early research [34]–[36], including investigating the effectiveness of deep education in the understanding of human activity, the related studies were carried out. In addition to the eventual creation of fundamental awareness of human activity, latest research is performed to face the unique challenges. Deep learning however, due to its sudden growth, busy progress, and lack of technical support, is facing resentful support by the researcher. It is also important to explain why deep learning in human activity is possible and effective given the difficulties.

- Deep learning is ”deep”, the most appealing attribute. Deep model layer by layer architectures make it possible to learn scalably from easy to abstract functionality. Advanced computing tools such as GPUs often allow deep models to learn descriptive functions from complex data. The outstanding ability to understand also helps the activitys identification system to closely evaluate multimodal sensory data and correctly identify them.
- Various neural network architectures represent multi-faceted functions. For example, convolutionary neural networks (CNNs) are able to capture multimodal sensory input locally and the local translation invariance is accurate [37]. Recurring neural networks (RNNs) remove temporal addiction and slowly acquire information over time to transmit sensory input to understand human activity.
- Deep neural networking can be detachable and scalable into interconnected networks with a global optimization feature that promotes various deep learning strategies learning [38], deep active education [39], a framework for deeper attention [40] and other approaches that are not systemic and effective [41], [42]. Works which take these techniques into account serve to numerous deep learning challenges.

C. Human Activity

Daily life activities (ADLs) are broadly defined. ADL's are all things we carry out every day, such as feeding, washing, dressing, working, homemaker, recreation and physical activity. The HAR scientific analysis offers an overview of the ADL's most researched.

Among the ADLs, walking, running, standing, sitting upstairs and walking downstairs are the most common activities in HAR study. However, in the past few years, other behaviours such as various phases of cooking, [43] house-cleansing, [44]–[46], smoking [47], swimming [48] etc., were also investigated. Other activities, including complex activities, were studied in recent years. Various experiments are conducted at certain sites like sitting on the floor, lying on the bed [49], walking and standing on elevator [50], walking and running on a treadmill, walking on a parking lot, stepping [51] or practising on a cross-trainer [51]. Additional comprehensive recognition of movement includes complex weapon movements such as transport/reaching of an object, its release, frontal height and other actions that can be carried out in connection with other objects [52], [53]. One main field of HAR research is population ageing and the increase in the number of persons with physical and cognitive impairments. Many HAR models are used to assist users in identifying and preventing risks such as falling in older adults in parkinson's disease [54]–[56] or freezing gait (foG) [24]. In addition, ADLs are becoming common for activity tracking devices. These devices can estimate physiological and physical parameters, for example, heart rate, blood pressure, steps, shift

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 June 2021 doi:10.20944/preprints202102.0349.v2

in levels and consumed calories. Advanced instruments can detect sleep and sleep neurological phases [57] (i.e. nREM and REM); all of the information processed can also be used as a HAR algorithm [57].

IV. Discussion & Challenges

We presented a summary of the latest HAR research in this paper. HAR is a crucial field in the identification of activity, computer technology and human support. In recent decades, HAR is becoming much more important with increasing new technologies and rising needs, including the ageing population. DL-based HAR methods have achieved outstanding recognition results in recent years. But CML-based methods are still commonly used and without any computational costs they produce excellent performance. The reproductivity of ML models has nevertheless become more and more relevant in recent years. According to our research the findings are not fully reproducible due to proprietary data sets for 78% of the proposed HAR methods. This leads to obstacles for the research community to identify and benchmark the best models.

In addition, the absence of heterogeneous public data sets decreases the opportunity to create HAR models with greater capacity for generalisation. The data used in the documents examined was mainly obtained in a regulated environment. The inter and intra-subject variability that is missing in such scription data sets compounds this issue because the majority of the proposed HAR models are only evaluated and collected in a single checked setting on a small number of activities. Of the 149 HAR models analysed, 87 were tested with a single dataset, while the other 62 were tested with several datasets. A further significant problem is the interpretability of the findings, mainly related to papers with similar methods and tests on the same dataset, which say that the accuracy of activity recognition results is almost identical. Such a problem involves testing using proprietary software, a shortage of open source code and writers who are not providing their source code publicly. Furthermore, the heterogeneity of the data and a HAR method concept, which recognises the behaviours of individuals with various motor and physical features, coincide with the data sources for data collection. A number of sensors and instruments are often used to collect data, as we have shown. In terms of data source, the methodologies suggested are nevertheless typically very rigid. Especially when using a certain sensor(s) and then changing a sensor model it becomes difficult for a specific person to test a methodology. Different sensors possess various technological features, which often include their particular state, e.g. measurement error or noise presented by a particular sensor. CML models are more common than complex DL-based models in the case of HAR models. The reason for this is that CML models need less training data and lower computer requirements. Furthermore, DL models are hard to interpret inherently. Nevertheless, DL models are exceptional in the knowledge and precision of more complex activities. Furthermore, a data preprocessing step is not needed. The selection of the exact DL or CML model was focused primarily on the computer specifications and the quantity of (marked) data available for training.

With recognition of human activity several forms of deep learning methodologies have been used. However, many technological challenges still face this area. Many of the problems are associated with other areas of pattern recognition, including machine vision and analysis of natural languages, while others are specific to sensor-based activity recognition. Below are few examples of issues that the recognition group will tackle.

- The first challenge is the question of extraction of features. The identification of operation is basically a classification concern, and it shares a similar difficulty with other classification problems, including the elimination of features. Feature selection is more difficult to identify sensor-based activity as there are differences in the inter-activities [31]. Related features of various activitys (for example, walking and running) can be noticed. Therefore, distinctive characteristics that reflect operations are difficult to create uniquely.
- Wide annotated data samples are needed for training and assessment of learning techniques. However, gathering and annotating sensory experience data is costly and time intensive. Annotation scarcity thus poses a major obstacle to understand sensor activity. Furthermore, it is especially difficult to collect data about any emergent or unpredictable events (e.g., accidental falling).
- Recognition of human activity comprises three elements: consumer, time and sensor. Second, the habits of activity depend on people. Different users may have different types of operation. Third, the definitions of operation differ over time. This is unworkable to conclude that consumers can remain static in their market habits for a long time. In addition, as modified, new activitys may occur. Fourth, numerous sensor systems are installed in human bodies or ecosystems on an opportunistic basis. Driven by events, the structure and configuration of sensors greatly affects results. These three allows the sensory input for action identification to be heterogeneous and desperately need to be mitigated.
- One factor that threatens understanding is the nature of the data connection. Data connection refers to the number of users and the number of operations for which data is associated. The identification of activity is driven by sophisticated data association and entails several individual challenges. Composite activitys are the first challenge. Many activities are focused on basic tasks, such as walking and sitting. Nonetheless, synthetic tasks that consist of a series of atomic events are more practical ways to document human everyday routines. For e.g., shooting on the tap, soaping, rubbing the hands, turning off the tap, ”washing hands” are provided. Data segmentation is one problem powered by composite operation. An activity composite can be described as an activity sequence. Accurate task identification thus depends heavily on specific methods of data segmentation. The third challenge is posed by overlapping events. Simultaneous events occur as the individual engages concurrently in multiple tasks such as listening to a phone call while watching TV. The scope

 doi:10.20944/preprints202102.0349.v2

of the data interaction is often related to the multi-occupant activitys. Recognition becomes difficult when a variety of individuals perform a series of actions, which typically occur in multi-resident situations.

- The reliability of the identification method of human activity is another consideration which needs to be concerned. Efforts must be taken to make the program accessible to a significant number of people, since knowledge of human activity can be multiplied in human everyday life. Next, the program would be usable to suit portable devices to provide an instant response. The problem of calculative costs will also be dealt with. Additionally, because the users' lives are constantly tracked by the identification program, there are chances of personal data leakage. Driving the device into private space is yet another matter that should be discussed.
- In comparison to photographs or text, sensory data is elusive and unreadable for action recognition. In addition, due to inherent sensor imperfections, sensory data invariably contains a lot of noise information. Therefore, accurate approaches for recognizing sensory data should be interpretable and able to recognize what aspect of the data makes identification simpler and which aspect can deteriorate that.

V. Conclusion & Future Direction

In the last decade, HAR systems have become an increasing field of research and have made remarkable progress. In particular, sensor-based HAR offers numerous advantages compared with vision-based HAR methods which are confidential and computational requirements limited. ML and DL-based activity identification algorithms become centralised in HAR. From a meta-examination of the current HAR surveys on the basis of the most studied human activities we analysed the examined literature, the most widely-used electronic data source sensors and most popular devices that integrate with these sensors without the video-based methodologies. Details were of great interest to sensor-based data from physiological, inertial and environmental sensors. Also thoroughly studied are the product categories categorised in: a) stand-alone devices, b) smartphones or c) smartwatch devices. The results were shown for each group based on the average number of activities recognised, the average number of data sets used to check the methods and the average accuracy. This study also examines accelerometer, gyroscope and magnetometer-based methodologies. We have addressed methods and findings on the basis of feature extraction, reduction of noise and standardisation. We provided a description of the most frequently used HAR recognition models. We have provided the most commonly used DL and ML models for this reason and their findings both in terms of consistency (exactitude) and quantity (number of recognised activities). We have concluded that HAR researchers still prefer traditional ML models, mainly because they need less data and lower computing power than DL models. However, many dynamic behaviours were more fully recognised by the DL models. Future studies should concentrate on methodology growth and identification of more complicated activities with more sophisticated generalisation capabilities.

A few potential directions for research are provided below, based on reviewed articles. The lack of structured methodologies that can lead to heterogeneous activities carried out by a variety of user is among the key shortcomings of HAR algorithms. In order to solve a similar problem, transfer learning may reuse the information that is gained in one problem. For example, information gained on the basis of a certain inertial sensor located at a specific position in the body can be theoretically reused with a different sensor location or a different inertial sensor type. The degree to which transfer learning, in different scenarios can be beneficial, is not fully explored and must be examined further. The fusion of sensors offers also a successful course. The fusion of different sensors, in particular, could address the reliability and accuracy of a single sensor and enrich the information collected. Where data is not accurate from a single mode, the device may switch to another sensor mode to ensure robust data collection. Fine grains of activity recognition based on regular object experiences are another direction of study. This will help us to identify sub-actions and action sequence and provide downstream applications with much richer background knowledge. Sensor fusion may also be useful when several inertial sensors or close-up sensors are connected to everyday objects. We include a collection of recommendations to further advance progress in this area. First, benchmark datasets for the HAR population should be developed as a priority. In order to demonstrate an upgrade, new HAR models should be compared with existing HAR models in reference results. In addition, it is highly recommended that you create data sets with a sufficient number of topics and a range of activities. Recognition of fine-grained activities can also benefit from the standardisation of benchmarks. In addition to designing and improving HAR algorithms, researchers working on HAR algorithm should pay attention to hardware and device issues. Calculation of devices and memory, CPU and battery use analysis should be the primary objective in examining the balance between the resource use and the accuracy of recognition. Lastly, a broad study of the dependent positions and orientation may lead to inconsistent and non-robust downstream implementations, if not the design of position-dependent techniques.

- Recognition of human activity requires sufficiently annotated samples to train the deep learning models. Unsupervised training can contribute to mitigating those needs. To date, deep unsupervised templates for identification of human activity are primarily used to distinguish characteristics but can not classify activitys as there are no ground-breaking facts. One possible approach to unsupervised training to infer true labels is therefore to look for additional information that leads to a common and deep unsupervised approach of transfer learning [58]. Another way forward is to use methods based on results, such as ontology [59].
- There is a huge struggle to identify new things that have never been used in models. A robust model would be able to acquire correct learning without any simple truth and learn new skills online. A good way is to learn functionality that

 doi:10.20944/preprints202102.0349.v2

can be used for different activities. Although [60] shows that mid-level characteristics can be used to represent activity with a variety of requirements, dissolute features [61] are another helpful approach for new activities.

- Potential anticipation of events expands the knowledge of activity. The activity prediction method, unlike activity detection, will forecast users' actions early. The predictive method is useful for human activity identification so that it can be used for intelligence systems, crime monitoring and driver behaviour. The actions are generally in a certain order in certain that activity tasks. Therefore, it is helpful to model the temporal dependency between events to predict future forecasts. For certain functions, LSTMs [62] are acceptable. LSTMs can not, however, incorporate these long-term dependencies for long-term operations. In this scenario, brain impulses [63] will help to encourage the estimation of activity.
- Although hundreds of works have been examined in deep learning and the perception of human activity by the sensors, state-of-the-art criteria for realistic comparisons are missing. The research conditions and assessment criteria for assessing activity detection efficiency differ from document to document. The separation of preparation / evaluation / testing affects the outcomes of identification, while deep learning is largely dependent on development evidence. This is also imperative that all studies have a mature standardization. It is worth noting that in many places such a issue is missing. To order to ensure fair contrast, ImageNet Challenge [64], for instance, specifics are clearly described. Jordao etal. [65] have carried out and tested a number of structured activities, but comprehensive and well-known standardization in the area of human activity identification is still not possible.
- With the growing elderly population, Ambient Assisted Living (AAL) has been thoroughly studied by Smart Home (SH) in order to promote independent livelihood. Recognition of Human Activity (HAR) is the cornerstone of AAL systems for the detection of daily life activity (DLA) and timely support. Existing SH based AAL systems are primarily focused on fine grained activity identification (AR) and are effective in using binary sensors for daily items. By checking the expected object interactions with evidence from many heterogeneous sensor data the fined-grain AR receives limited attention.
- In the context-aware life assistive systems, activity recognition (AR) is central. The segmentation of sensor events observed when interleaved or concurrent everyday life activities (DDAs) take place in AR is a challenge. Several studies have proposed methods to separate and organise sensor observations and to recognise the basic or composite performance of generic ADLs. In semantically differentiating individual sensor events, however, little has been explored [15].
- Human Activity Recognition (HAR) systems with wearables usually rely on datasets manually annotated by human experts for precise timing of cases of relevant activity. However, in the mostly mobile scenarios of recognition of human activity, it is often very difficult to obtain data annotations like this. This also results in a degree of ambiguity on labels.This area can be explored more extensively [16].

REFERENCES

[1] Rodolfo S Antunes, Lucas A Seewald, Vinicius F Rodrigues, Cristiano A Da Costa, Luiz Gonzaga Jr, Rodrigo R Righi, Andreas Maier, Bjoern Eskofier, Malte Ollenschlaeger, Farzad Naderi, et al. A survey of sensors in healthcare workflow monitoring. *ACM Computing Surveys (CSUR)*, 51(2):1–37, 2018.

[2] Florenc Demrozi, Graziano Pravadelli, Azra Bihorac, and Parisa Rashidi. Human activity recognition using inertial, physiological and environmental sensors: A comprehensive survey. *IEEE Access*, 8:210816–210836, 2020.

[3] Yan Wang, Shuang Cang, and Hongnian Yu. A survey on wearable sensor modality centred human activity recognition in health care. *Expert Systems with Applications*, 137:167–190, 2019.

[4] Oscar D Lara and Miguel A Labrador. A survey on human activity recognition using wearable sensors. *IEEE communications surveys & tutorials*, 15(3):1192–1209, 2012.

[5] Ye Liu, Liqiang Nie, Li Liu, and David S Rosenblum. From action to activity: sensor-based activity recognition. *Neurocomputing*, 181:108–115, 2016.

[6] Ming Zeng, Le T Nguyen, Bo Yu, Ole J Mengshoel, Jiang Zhu, Pang Wu, and Joy Zhang. Convolutional neural networks for human activity recognition using mobile sensors. In *6th International Conference on Mobile Computing, Applications and Services*, pages 197–205. IEEE, 2014.

[7] Allan Stisen, Henrik Blunck, Sourav Bhattacharya, Thor Siiger Prentow, Mikkel Baun Kjærgaard, Anind Dey, Tobias Sonne, and Mads Møller Jensen. Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition. In *Proceedings of the 13th ACM conference on embedded networked sensor systems*, pages 127–140, 2015.

[8] Eiman Kanjo, Eman MG Younis, and Chee Siang Ang. Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection. *Information Fusion*, 49:46–56, 2019.

[9] Natalia Neverova, Christian Wolf, Griffin Lacey, Lex Fridman, Deepak Chandra, Brandon Barbello, and Graham Taylor. Learning human identity from motion patterns. *IEEE Access*, 4:1810–1820, 2016.

[10] Li Liu, Yuxin Peng, Ming Liu, and Zigang Huang. Sensor-based human activity recognition system with a multilayered model using time series shapelets. *Knowledge-Based Systems*, 90:138–152, 2015.

[11] Jeffrey Donahue, Lisa Anne Hendricks, Sergio Guadarrama, Marcus Rohrbach, Subhashini Venugopalan, Kate Saenko, and Trevor Darrell. Long-term recurrent convolutional networks for visual recognition and description. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2625–2634, 2015.

[12] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. nature 521 (7553), 436-444. *Google Scholar Google Scholar Cross Ref Cross Ref*, 2015.

[13] Jindong Wang, Yiqiang Chen, Shuji Hao, Xiaohui Peng, and Lisha Hu. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 119:3–11, 2019.

[14] Reda Elbasiony and Walid Gomaa. A survey on human activity recognition based on temporal signals of portable inertial sensors. In *International Conference on Advanced Machine Learning Technologies and Applications*, pages 734–745. Springer, 2019.

[15] D. Triboan, L. Chen, and F. Chen. Fuzzy-based fine-grained human activity recognition within smart environments. In *2019 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pages 94–101, 2019.

[16] Darpan Triboan, Liming Chen, Feng Chen, and Zumin Wang. A semantics-based approach to sensor data segmentation in real-time activity recognition. *Future Generation Computer Systems*, 93:224–236, 2019.

 doi:10.20944/preprints202102.0349.v2

[17] Henry Friday Nweke, Ying Wah Teh, Ghulam Mujtaba, and Mohammed Ali Al-Garadi. Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions. *Information Fusion*, 46:147–170, 2019.

[18] Oliver Faust, Yuki Hagiwara, Tan Jen Hong, Oh Shu Lih, and U Rajendra Acharya. Deep learning for healthcare applications based on physiological signals: A review. *Computer methods and programs in biomedicine*, 161:1–13, 2018.

[19] Sreenivasan Ramasamy Ramamurthy and Nirmalya Roy. Recent trends in machine learning for human activity recognition—a survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4):e1254, 2018.

[20] Jafet Morales and David Akopian. Physical activity recognition by smartphones, a survey. *Biocybernetics and Biomedical Engineering*, 37(3):388–400, 2017.

[21] L. Chen, C. D. Nugent, and H. Wang. A knowledge-driven approach to activity recognition in smart homes. *IEEE Transactions on Knowledge and Data Engineering*, 24(6):961–974, 2012.

[22] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu. Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6):790–808, 2012.

[23] L. Chen, C. Nugent, and G. Okeyo. An ontology-based hybrid approach to activity modeling for smart homes. *IEEE Transactions on Human-Machine Systems*, 44(1):92–105, 2014.

[24] Florenc Demrozi, Ruggero Bacchin, Stefano Tamburin, Marco Cristani, and Graziano Pravadelli. Toward a wearable system for predicting freezing of gait in people affected by parkinson's disease. *IEEE journal of biomedical and health informatics*, 24(9):2444–2451, 2019.

[25] Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006.

[26] Pedro Domingos. A few useful things to know about machine learning. *Communications of the ACM*, 55(10):78–87, 2012.

[27] Chelsea Dobbins and Reza Rawassizadeh. Towards clustering of mobile and smartwatch accelerometer data for physical activity recognition. In *Informatics*, volume 5, page 29. Multidisciplinary Digital Publishing Institute, 2018.

[28] Amari Vaughn, Paul Biocco, Yang Liu, and Mohd Anwar. Activity detection and analysis using smartphone sensors. In *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, pages 102–107. IEEE, 2018.

[29] Zahraa Said Abdallah, Mohamed Medhat Gaber, Bala Srinivasan, and Shonali Krishnaswamy. Adaptive mobile activity recognition system with evolving data streams. *Neurocomputing*, 150:304–317, 2015.

[30] Tâm Huynh and Bernt Schiele. Analyzing features for activity recognition. In *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies*, pages 159–163, 2005.

[31] Andreas Bulling, Ulf Blanke, and Bernt Schiele. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR)*, 46(3):1–33, 2014.

[32] Jessica Lin, Eamonn Keogh, Stefano Lonardi, and Bill Chiu. A symbolic representation of time series, with implications for streaming algorithms. In *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, pages 2–11, 2003.

[33] Samira Pouyanfar, Saad Sadiq, Yilin Yan, Haiman Tian, Yudong Tao, Maria Presa Reyes, Mei-Ling Shyu, Shu-Ching Chen, and SS Iyengar. A survey on deep learning: Algorithms, techniques, and applications. *ACM Computing Surveys (CSUR)*, 51(5):1–36, 2018.

[34] Sojeong Ha, Jeong-Min Yun, and Seungjin Choi. Multi-modal convolutional neural networks for activity recognition. In *2015 IEEE International conference on systems, man, and cybernetics*, pages 3017–3022. IEEE, 2015.

[35] Nicholas D Lane and Petko Georgiev. Can deep learning revolutionize mobile sensing? In *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications*, pages 117–122, 2015.

[36] Jianbo Yang, Minh Nhut Nguyen, Phyo Phyo San, Xiao Li Li, and Shonali Krishnaswamy. Deep convolutional neural networks on multichannel time series for human activity recognition. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

[37] Nils Y Hammerla, Shane Halloran, and Thomas Plötz. Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv preprint arXiv:1604.08880*, 2016.

[38] Ali Akbari and Roozbeh Jafari. Transferring activity recognition models for new wearable sensors with deep generative domain adaptation. In *Proceedings of the 18th International Conference on Information Processing in Sensor Networks*, pages 85–96, 2019.

[39] Gautham Krishna Gudur, Prahalathan Sundaramoorthy, and Venkatesh Umaashankar. Activeharnet: Towards on-device deep bayesian active learning for human activity recognition. In *The 3rd International Workshop on Deep Learning for Mobile Systems and Applications*, pages 7–12, 2019.

[40] Vishvak S Murahari and Thomas Plötz. On attention models for human activity recognition. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers*, pages 100–103, 2018.

[41] Chihiro Ito, Xin Cao, Masaki Shuzo, and Eisaku Maeda. Application of cnn for human activity recognition with fft spectrogram of acceleration and gyro sensors. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, pages 1503–1510, 2018.

[42] Akhil Mathur, Tianlin Zhang, Sourav Bhattacharya, Petar Velickovic, Leonid Joffe, Nicholas D Lane, Fahim Kawsar, and Pietro Lió. Using deep data augmentation training to address software and hardware heterogeneities in wearable and smartphone sensing devices. In *2018 17th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 200–211. IEEE, 2018.

[43] Ferhat Attal, Samer Mohammed, Mariam Dedabrishvili, Faicel Chamroukhi, Latifa Oukhellou, and Yacine Amirat. Physical human activity recognition using wearable sensors. *Sensors*, 15(12):31314–31338, 2015.

[44] Mingqi Lv, Ling Chen, Tieming Chen, and Gencai Chen. Bi-view semi-supervised learning based semantic human activity recognition using accelerometers. *IEEE Transactions on Mobile Computing*, 17(9):1991–2001, 2018.

[45] Ye Liu, Liqiang Nie, Lei Han, Luming Zhang, and David S Rosenblum. Action2activity: recognizing complex activities from sensor data. *arXiv preprint arXiv:1611.01872*, 2016.

[46] Muhammad Arif and Ahmed Kattan. Physical activities monitoring using wearable acceleration sensors attached to the body. *PloS one*, 10(7):e0130851, 2015.

[47] Edwin Valarezo Añazco, Patricio Rivera Lopez, Sangmin Lee, Kyungmin Byun, and Tae-Seong Kim. Smoking activity recognition using a single wrist imu and deep learning light. In *Proceedings of the 2nd international conference on digital signal processing*, pages 48–51, 2018.

[48] Gino Brunner, Darya Melnyk, Birkir Sigfússon, and Roger Wattenhofer. Swimming style recognition and lap counting using a smartwatch and deep learning. In *Proceedings of the 23rd International Symposium on Wearable Computers*, pages 23–31, 2019.

[49] Donghui Wu, Zhelong Wang, Ye Chen, and Hongyu Zhao. Mixed-kernel based weighted extreme learning machine for inertial sensor based human activity recognition with imbalanced dataset. *Neurocomputing*, 190:35–49, 2016.

[50] Gabriele Civitarese, Riccardo Presotto, and Claudio Bettini. Context-driven active and incremental activity recognition. *arXiv preprint arXiv:1906.03033*, 2019.

[51] Lukun Wang. Recognition of human activities using continuous autoencoders with wearable sensors. *Sensors*, 16(2):189, 2016.

[52] Abdulhamit Subasi, Dalia H Dammas, Rahaf D Alghamdi, Raghad A Makawi, Eman A Albiety, Tayeb Brahimi, and Akila Sarirete. Sensor based human activity recognition using adaboost ensemble classifier. *Procedia Computer Science*, 140:104–111, 2018.

[53] Abdulhamit Subasi, Mariam Radhwan, Rabea Kurdi, and Kholoud Khateeb. Iot based mobile healthcare system for human activity recognition. In *2018 15th learning and technology conference (L&T)*, pages 29–34. IEEE, 2018.

[54] Genming Ding, Jun Tian, Jinsong Wu, Qian Zhao, and Lili Xie. Energy efficient human activity recognition using wearable sensors. In *2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pages 379–383. IEEE, 2018.

[55] Daniela Micucci, Marco Mobilio, and Paolo Napoletano. Unimib shar: A dataset for human activity recognition using acceleration data from smartphones. *Applied Sciences*, 7(10):1101, 2017.

 doi:10.20944/preprints202102.0349.v2

[56] Yiming Tian, Xitai Wang, Wei Chen, Zuojun Liu, and Lifeng Li. Adaptive multiple classifiers fusion for inertial sensor based human activity recognition. *Cluster Computing*, 22(4):8141–8154, 2019.
[57] Girija Chetty, Matthew White, and Farnaz Akther. Smart phone based data mining for human activity recognition. *Procedia Computer Science*, 46:1181–1187, 2015.
[58] Yoshua Bengio. Deep learning of representations for unsupervised and transfer learning. In *Proceedings of ICML workshop on unsupervised and transfer learning*, pages 17–36, 2012.
[59] Daniele Riboni, Linda Pareschi, Laura Radaelli, and Claudio Bettini. Is ontology-based activity recognition really effective? In *2011 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pages 427–431. IEEE, 2011.
[60] Harideep Nair, Cathy Tan, Ming Zeng, Ole J Mengshoel, and John Paul Shen. Attrinet: learning mid-level features for human activity recognition with deep belief networks. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, pages 510–517, 2019.
[61] Luan Tran, Xi Yin, and Xiaoming Liu. Disentangled representation learning gan for pose-invariant face recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1415–1424, 2017.
[62] Moez Baccouche, Franck Mamalet, Christian Wolf, Christophe Garcia, and Atilla Baskurt. Action classification in soccer videos with long short-term memory recurrent neural networks. In *International Conference on Artificial Neural Networks*, pages 154–159. Springer, 2010.
[63] Dalin Zhang, Lina Yao, Xiang Zhang, Sen Wang, Weitong Chen, Robert Boots, and Boualem Benatallah. Cascade and parallel convolutional recurrent neural networks on eeg-based intention recognition for brain computer interface. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.
[64] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252, 2015.
[65] Artur Jordao, Antonio C Nazare Jr, Jessica Sena, and William Robson Schwartz. Human activity recognition based on wearable sensor data: A standardization of the state-of-the-art. *arXiv preprint arXiv:1806.05226*, 2018.