Effectiveness of Machine Learning on Human Activity Recognition Using Accelerometer and Gyroscope Sensors: A Survey

Munid ALANAZI
Department of Electrical Engineering and Computer Science
Florida Atlantic University
Boca Raton, FL, USA

Raghdah SAEM ALDAHR
Department of Electrical Engineering and Computer Science
Florida Atlantic University
Boca Raton, FL, USA

Mohammad ILYAS
Department of Electrical Engineering and Computer Science
Florida Atlantic University
Boca Raton, FL, USA

ABSTRACT

HAR is defined as using machine learning to classify certain human activities in specific time periods by learning from inertial sensor time series data [2]. Over the past few years, the growth in the computational field has been robust enough to transfer a world to a more intelligent place. Currently, the electronic parts turn out to be available in many shapes and sizes. For example, accelerometer and gyroscope sensors could be manufactured into a single piece that can be operated in wired or wireless settings (Bluetooth). Also, this piece could be used in smartphones because the smartphone cannot operate most of its features without these two sensors with respect to the other sensor (magnetometer and inclinometer). The amount of data captured from these sensors in a time series format is in billions of bytes. Human activity recognition is one of the important applications that could be implemented by using these sensors blended with certain machine learning algorithms. Such applications have become important research area because these serve athletic, healthcare, and personal use. This paper summarizes the important aspects of machine learning, human activity recognition, and reviews existing scientific literature in the field of human activity recognition.

Keywords: Human Activity Recognition, Accelerometer, Gyroscope, Sensors, Machine Learning

1. INTRODUCTION

HAR is defined as using machine learning to classify certain human activities in specific time periods by learning from inertial sensor time series data [1].The field of Human Activity Recognition (HAR) has gained increasing interest by researchers since it affects many, such as those engaged in athletic activities, as well as employees in the healthcare industry. On one hand, the advancements of the sensors and how they communicate with each other and with the actuators (sensor networks) helps to bring forth robust applications in various fields [2]. On the other hand, the sensors come in multiple shapes, such as internal sensors (smartphone sensors) and external sensors. Based on these sensors, tri-axial data can be captured from them. These data can help to recognize physical object activities by implementing machine learning algorithms. Machine learning techniques, such as Random Forest (RF), support vector machine (SVM), and additional algorithms have been used in HAR based on smart phone data. In addition, Deep Learning has been used in HAR also. Convolutional Neural Networks and Recurrent Neural Networks have been applied on some datasets in the field of Activity Recognition and many other fields [3]. There are a couple of activities that have been classified by some other researchers' models shown in Table 1.

Dataset	Some of the classified activities	Reference
KU-HAR	STAND, SIT, TALK-SIT, TALK STAND, STAND-SIT	4
UCI	WALKING, WALKING_DOWNSTAIRS , WALKING_UPSTAIRS, LAYING SITTING, STANDING	5
WISDM	WALKING, JOGGING, SITTING, STANDING	6

Table 1: Activity classes in previous research.

The rest of the paper is organized as follows. Section 2 provides an overview of machine learning and how it works with human activity recognition. Section 3 gives insights on human activity recognition with important facts and figures. Section 4 reviews the latest papers found in literature which performed classification approaches for human actions using machine learning and deep learning algorithms. Section 5 concludes the paper and mentions some future directions.

2. MACHINE LEARNING

Machine learning has powerful abilities that allow its algorithms to accurately predict some complex patterns of data [7]. The use of machine learning in different fields has proven its strength in solving some of the problems and challenges. With the advancement of data science and analytics, machine learning became a significant support to such large amounts of data. Every

ISBN: 978-1-950492-65-7
ISSN: 2771-0947

https://doi.org/10.54808/WMSCI2022.02.1

machine learning model has a specific pipeline which explains the process of the model from the data collection to model deployment/visualization. The model pipeline may differ from model to another, depending on the required processes. In HAR, the model pipeline as shown in Figure 1, has a specific structure as the HAR application deals with signal data which has some noise that needs to be cleaned with some preprocessing techniques.

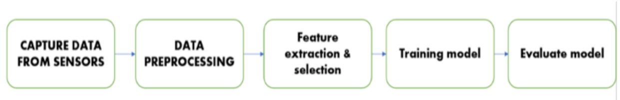

Figure 1: Machine learning pipeline used in Human activity recognition.

The classification of binary or multiclass is one of the concepts in machine learning which is used in different domains and applications. Therefore, the huge amount of data leads the machine learning features to a specific class. The power of the algorithms in machine learning can help to classify the data to many classes used in a specific dataset. Mostly, the machine learning models evaluated in terms of recall, precision, F-measure, and accuracy [8]. The accuracy shows how accurate the model is in classifying the defined classes in the dataset. For example, Human Activity dataset may have some classes, such as standing, walking, climbing stairs, etc. The model should have high accuracy in predicting these classes and decreasing the misclassified classes, to have an efficient HAR model. The other metrics could also help to prove the efficiency of the model.

Machine learning could be applied through multiple tools where some of them do not need coding skills, but further know how to build a model in the correct way. Python is used to implement the machine learning because it has some predefined libraries that are used to build models which are: mlpy, tensorflow, and PyTourch. Each one can perform specific tasks related to machine learning. Machine learning python (mlpy) provides some machine learning methods to solve supervised and unsupervised problems. It is an open-source library built on NumPy, SciPy, and GNU scientific libraries [9]. Some libraries details are shown in Table 2.

Library Name	Written in	Developer	Release
NumPY	Python, C	Community project	2006
SciPy	Python, Fortran, C, C++	Community library project	2001
PyTourch	Python; C++; CUDA	Facebook's AI Research lab (FAIR)	2016

Table 2: Types of signals and uses.

As previously mentioned, some of the tools can be applied without a need for normal coding in Python or any other programming language. The Waikato Environment for Knowledge Analysis (WEKA) is one of the tools that is extensively used by many researchers, that builds machine learning models. It consists of data preprocessing tools and ensembles machine learning algorithms which help people from academia and business to implement machine learning [10].

3. HUMAN ACTIVITY RECOGNITION SYSTEM

In 2018, the number of smart phone users in the world was 2.53 billion. After two years, in 2020, this number increased to 3.5 billion. In 2021, this number is expected to reach 3.8 billion [11]. Consequently, the huge number of smart phone users will increase every year and that will help the machine learning engineers to utilize the smart phone hardware elements such as accelerometer and gyroscope sensors to build human activity recognition system that would be beneficial for some fields, like healthcare.

HAR is the concept of recognizing the daily activities of humans based on some captured signals. Because of the various areas of applications in HAR, the research work of HAR has become important and significant in the past few years. Wearable sensors are also used to detect human activities.

There are categories of wearable sensor signals where some of them utilize the HAR system [8]. In Table 3, signal categories are organized into two columns: the first column lists the type of signals, and the second column lists the use of each signal [8].

Signal Name	Use
Physiological signals	Heart Rate, Skin temperature, electrocardiography
Environmental signals	Humdity, audio level, temperature
Location data	Globa positioning system (GPS), intergrated GPS
Acceleration signals	Body Movement Recording

Table 3: Types of signals and uses.

Usually, the signal data has some noise because of some known reasons that happen during the data capturing process. This noise can be cleaned using some filtering techniques such as, Median filter and bandpass Butterworth filter. Also, in most of the cases, the sliding window technique can be used to segment the data. After that, feature extraction is applied to the data for both time and frequency domains. In some cases, only one of them is applied. At this time, the data become ready to be trained in the model. There are a lot of classifiers that have been used in HAR. For example, SVM, Naive Bayes, Decision Tree, Random Forest, etc. Finally, the classification performance measured using some common metrics used in most of machine learning models are accuracy, precision, recall, F-measure, and operating characteristic (ROC).

Most of the performance evaluations used in research studies are the four common expressions we mentioned before[12]. The expressions are accuracy, precision, recall, F-measure and written as below:

Metric	Mathmatical expression	Acronyms index
Accuracy	$Acc = \frac{TP + TN}{TP + TN + FP + FN}$	TN: TRUE NEGATIVE TP: TRUE POSITIVE FN: FALSE NEGATIVE FP: FALSE POSITIVE
Recall	$Rc = \frac{TP}{TP + FN}$	
Precision	$\Pr = \frac{TP}{TP + FP}$	
F1-Score	$F1s = \frac{2 \times \Pr \times Rc}{\Pr + Rc}$	

Table 4: Performance evaluation metrics [13].

There are additional metrics that have been used extensively by Machine and Deep Learning researchers, which is a table that presents the results of prediction in classification problems. The table below shows an example of confusion matrix for two classes in classification problems.

Type of class	Predicted class	
ACTUAL CLASS	CLASSIFIED	MISCLASSIFIED
CLASSIFIED	True positive (TP)	False positive (FP)
MISCLASSIFED	False negative(FN)	True negative(TN)

Table 5: Confusion matrix for two classes [14].

4. LITERATURE REVIEW

HAR has been researched by many researchers who propose multiple approaches to solve the problem of activity recognition in the recent years. We focused on the researchers who have used inertial sensors data on their research. Some of them used machine learning and deep learning approaches on their classification problems. There are some high accuracy detection systems that have been found recently by using different approaches. In this survey paper, we have searched through Saudi Digital Library the English knowledge resources which it gives us a middleware access to the well-known databases such as Scopus, Web of Science, and IEEE, etc. we have read 25 papers and consider 10 for comprehensive review where they are related to HAR and use Smart phone inertial sensor.

In [15], the authors present some techniques for processing and training HAR dataset that are recorded from smartphone sensors. So, the detection system for activity recognition could have high accuracy. Also, they list the most common algorithms used in the field of human activity recognition which are Support Vector Machines (SVMs), k-Nearest Neighbors (KNN), Gaussian Mixture Models (GMM), Random Forests (RF), Markov Chains, and Hidden Markov Models (HMM). The feasibility of activity recognition was extensively researched in this paper that helped the researchers to find new opportunities to work on in the future.

In [16], a model proposed by authors that is based on a Convolutional Long Short-Term Memory (ConvLSTM) network, used the deep learning approach to classify some activities performed by humans using inertial smartphone sensors within the boundaries of indoor localization. The model succeeds in recognizing nine activities which are walking, going up in an elevator, not moving, going down in an elevator, walking downstairs, running, walking upstairs, or going up and down a ramp. In addition to this result, the human activities data could help in recognizing if someone has changed from one floor to another in the same building. Also, the error rate of localization system could be reduced with the smartphone activity data. In summary, authors found from their experiment that the human activity information improves the performance of indoor localization.

In [17], the Deep Neural Network is applied to recognize some human activities using a public dataset. The data is captured from inertial sensors which are accelerometer and gyroscope. Six activities have been recognized and its model performance is compared to some other classical machine learning algorithms. The accuracy of recognition for this approach has reached 98%, which is considered higher than most researchers in the field.

In [18], the authors present the significance of smartphone sensors used in human activity recognition because of the low cost and the size which is very important in some of HAR applications. Some feature selection and extraction have been researched and compared in terms of efficiency. This work resulted in some important notes such as the comparison between dynamic and static feature efficiency, where authors found dynamic features are more efficient. Also, some deep learning approaches have been tested to figure out which classifiers are better for this kind of application. They found that Conventional Neural Networks better than Support Vector Machine (SVM), K-Nearest Neighbor (KNN), or Multilayer Perceptron (MLP).

In [19], fall detection systems are presented by the authors by testing some enhanced features selection and classification methods. They used a public benchmark dataset, which has 12 activities and 4 kinds of falls by 66 subjects. This data has been tested using two different approaches. First, the detection of 6 activities has 99.9% accuracy. Second, the detection of all activities and falls together reached 96.8% accuracy.

In [20], a smartphone inertial sensors-based approach was presented for recognizing human activity. Their work mainly aimed at developing a robust human activity recognition system that is based on the data of the smartphone sensors. Smartphones appear to be very appropriate to use for activity recognition purposes since these devices are mainly and widely used by people in their daily lives for varied applications that include communication, healthcare, and others. Therefore, the authors of this paper proposed a new approach for activity recognition through the usage of smartphone inertial sensors like gyroscope sensors and accelerometers. Various robust features have been extracted from the sensor signals followed by kernel principal component analysis (KPCA) for the purpose of dimension reduction. Moreover, for activity training and recognition, the authors combined the robust features with the Deep Learning technique, Deep Belief Network (DBN). The proposed technique outperformed the traditional expression recognition techniques like Artificial Neural Network (ANN) and typical multiclass Support Vector Machine (SVM). Twelve different physical activities of the systems were examined, where the mean recognition rate was found to be 89.61% while the overall accuracy was found to be 95.85%. Conversely, the best results that the other traditional techniques may achieve, were a mean recognition rate of 82.02% and an overall accuracy rate of 94.12%. it was also clear that the system was capable of differentiating between basic transitional and non-transitional activities.

In their paper [21], the authors presented a deep convolutional neural network (convnet) to carry out effective and efficient HAR using the sensors of a smartphone through utilizing the 1D time-series signals and the inherent characteristics of activities. Simultaneously, they found a method to extract robust features from raw data automatically and data-adaptively. As proven by experiments, relevant and more complex features are actually extracted by convnets with each additional layer, despite the fact that the difference of the level of feature complexity decreases with each layer. A broader time span of temporal local correlation may be utilized (1×9–1×14) and a small pooling size (1×2–1×3) is proven to be useful. In addition, on moving activities, convnets attained an almost ideal classification, particularly very identical activities which were formerly identified to be very hard to classify. It was shown that the convnet performance easily outperforms the other state-of-the-art data mining approaches in HAR, where convnet technique achieved an accuracy of 95.75% with additional data of Fast Fourier Transform from HAR data set, and an accuracy of 94.79% with raw sensor data. Mostly, the reason of achieving a high accuracy rate is the almost ideal moving activities classification, particularly very identical activities like walking downstairs and walking upstairs, which were formerly identified to be very difficult to distinguish. Nevertheless, SVM outperformed convent's confusion in the classification of stationary activities.

In their paper [22], the authors presented an activity recognition system with the objective of detecting the activities through utilizing a set of classifiers approaches using the Wireless Sensor Data Mining (WISDM). Their system recognized six different activities, specifically jogging, standing, walking, sitting, downstairs and upstairs. For the purpose of determining the preferable classifier combination for activity recognition, many experiments were carried out. The experiments showed that the performance was better when the classifiers were combined. The authors used AdaBoost in combination with decision tree algorithm C4.5 to build an ensemble system. The system efficiently improved the performance with an accuracy level of 94.04%.

In [23], the research work aimed at studying the role of gyroscope sensors, accelerometers, and their combination in detecting, analyzing, and recognizing automatic human activity through the usage of artificial neural networks. Their experiment showed that accelerometer and gyroscope sensors may be used individually to recognize human activities, and that better performance was observed when both sensors were combined. Nevertheless, the usage of multiple sensors may create significant challenges because of the low battery capacity of the mobile phones, or in other words, the mobile phone battery limitations. Continuous sensing by the mobile phones is required for activity recognition. The experimental results found on the published dataset show that it is possible to use each one of the sensors separately for human activity recognition. Nevertheless, the results showed that the data of the accelerometer had better performance than the data of the gyroscope sensor with a classification accuracy of 92%. In addition, the performance was improved when gyroscope sensors and accelerometers were combined than when they were used individually with an accuracy rate of 95%.

In [24], the authors suggested a model for recognizing the human physical activity based on the collected data from the sensors of smartphones. The suggested technique implies the usage of three sensors available on a smartphone to develop a classifier. Those sensors are gyroscope, gravity sensor and accelerometer. The reason that made the authors choose to carry out their solution on mobile phones was the fact that they are widely used, and that no subjects are required to carry extra sensors which may delay their activities. The human activities that they targeted for their proposal were as follows: running, walking, standing, sitting, descending stairs and ascending stairs. The authors assessed the solution against a total of two datasets as follows: one internal dataset collected by them and one external dataset, with significant effect. The results revealed good accuracy for recognizing all the six activities especially for running, standing, walking and sitting. Their model was fully executed on a mobile device as an application of Android. Based on the evaluation of results, most of the activities were recognized in a correct way, where four of those activities averaged an accuracy of 93%. Regarding the remaining two activities, even though lower scores were obtained for them, they scored a minimum accuracy of 86%. It can be seen from these numbers that utilizing sensors for user activity recognition is more reliable than ever. In addition, the suggested solution was also examined against an external standard dataset, where the results showed a little decrease in accuracy to 77%. Nevertheless, since the external dataset included data from various sensors, the results might have been affected. An Android application was used to materialize the paper which facilitated its usage by any user, even older adults.

Algorithms	**Activities**	**Types of data**	**accuracy**	**Ref**
(SVMs) (KNN) (GMM), (RF), Markov Chains, and (HMM).	6	Smart Phone sensor data	75%-90%	15
ConvLSTM	9	Smart Phone sensor data	73%	16
Deep neural network	6	Smart Phone sensor data	98%	17
(CNN), (SVMs), (KNN), (MLP)	6	Smart Phone sensor data	SVM: 96% CNN:98%	18
IBk, J48 decision tree, logistic regression, MLP	12	Smart Phone sensor data	98.8%	19
Deep Belief Network (DBN)	12	Smart Phone sensor data	95.85%	20
CNN	6	Smart Phone sensor data	95.75%	21
AdaBoost & C.45	6	Smart Phone sensor data	94.04%	22
MLP	6	Smart Phone sensor data	95%	23
MLP	6	Smart Phone sensor data	93% & 86%	24

Table 6: Comparison between literature algorithms and results.

5. CONCLUSION

Nowadays, artificial intelligence and machine learning models development have attracted the attention of an increasing number of researchers. HAR is one application that is concerned with identifying the humans' physical activities by implementing machine learning algorithms. It has many significant applications used in different sectors such as military and healthcare. Therefore, solving HAR related problem would significantly help a lot of professionals in other fields. Researchers have found several efficient models that could be deployed in real products with effective responses. In this paper we provided an overview of machine learning, human activity recognition systems, and a review of available literature related to activity recognition applications. For future research, one of our goals is perform a classification approach using Machine Learning and Deep Learning on the recent HAR dataset collected from smartphones and try to build accurate models that contribute to this promising field.

6. REFERENCES

[1] Raza, A., Tran, K. P., Koehl, L., Li, S., Zeng, X., & Benzaidi, K. (2021). Lightweight Transformer in Federated Setting for Human Activity Recognition. arXiv preprint arXiv:2110.00244.

[2] Bengalur, M. D. (2013, August). Human activity recognition using body pose features and support vector machine. In 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 1970-1975). IEEE.

[3] Abidalkareem, A. J., Abd, M. A., Ibrahim, A. K., Zhuang, H., Altaher, A. S., & Ali, A. M. (2020, July). Diabetic retinopathy (DR) severity level classification using multimodel convolutional neural networks. In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 1404-1407). IEEE.

[4] Sikder, N., & Nahid, A. A. (2021). KU-HAR: An open dataset for heterogeneous human activity recognition. Pattern Recognition Letters, 146, 46-54.

[5] Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. A Public Domain Dataset for Human Activity Recognition Using Smartphones. 21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013. Bruges, Belgium 24-26 April 2013.

[6] Jennifer R. Kwapisz, Gary M. Weiss and Samuel A. Moore (2010). Activity Recognition using Cell Phone Accelerometers, Proceedings of the Fourth International Workshop on Knowledge Discovery from Sensor Data (at KDD-10), Washington DC.

[7] Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Interpretable machine learning: definitions, methods, and applications. arXiv preprint arXiv:1901.04592.

[8] Elamvazuthi, I., Izhar, L. I., & Capi, G. (2018). Classification of human daily activities using ensemble methods based on smartphone inertial sensors. Sensors, 18(12), 4132.

[9] Albanese, D., Visintainer, R., Merler, S., Riccadonna, S., Jurman, G., & Furlanello, C. (2012). mlpy: Machine learning python. arXiv preprint arXiv:1202.6548..

[10] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. ACM SIGKDD explorations newsletter, 11(1), 10-18.

[11] D. Georgiev, 60 + Revealing Statistics about Smartphone Usage in 2020, (2020). https://techjury.net/blog/smartphone-usage- statistics/ (accessed May 27, 2020).

[12] Altaher, A. S., & Taha, S. M. R. (2017). Personal authentication based on finger knuckle print using quantum computing. International Journal of Biometrics, 9(2), 129-142.

[13] Alrowaily, M., Alenezi, F., & Lu, Z. (2019, July). Effectiveness of machine learning based intrusion detection systems. In International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage (pp. 277-288). Springer, Cham.

[14] Ting K.M. (2017) Confusion Matrix. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning and Data Mining. Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-7687-1_50

[15] Ilisei, D., & Suciu, D. M. (2019, October). Human-activity recognition with smartphone sensors. In OTM Confederated International Conferences" On the Move to Meaningful Internet Systems" (pp. 179-188). Springer, Cham.

[16] Moreira, D., Barandas, M., Rocha, T., Alves, P., Santos, R., Leonardo, R., ... & Gamboa, H. (2021). Human Activity Recognition for Indoor Localization Using Smartphone Inertial Sensors. Sensors, 21(18), 6316.

[17] Chen, Y., & Zhong, K. (2015). Human Activity Recognition Using Smartphone Sensor Data Via Deep Neural Networks.

[18] Nakano, K., & Chakraborty, B. (2017, November). Effect of dynamic feature for human activity recognition using smartphone sensors. In 2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST) (pp. 539-543). IEEE.

[19] Chatzaki, C., Pediaditis, M., Vavoulas, G., & Tsiknakis, M. (2016, April). Human daily activity and fall recognition using a smartphone's acceleration sensor. In International Conference on Information and Communication Technologies for Ageing Well and e-Health (pp. 100-118). Springer, Cham.

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

[21] 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.

[22] Daghistani, T., & Alshammari, R. (2016). Improving accelerometer-based activity recognition by using ensemble of classifiers. Int. J. Adv. Comput. Sci. Appl, 7(5), 128-133.

[23] Ogbuabor, G., & La, R. (2018, February). Human activity recognition for healthcare using smartphones. In Proceedings of the 2018 10th international conference on machine learning and computing (pp. 41-46).

[24] Voicu, R. A., Dobre, C., Bajenaru, L., & Ciobanu, R. I. (2019). Human physical activity recognition using smartphone sensors. Sensors, 19(3), 458.