WORLD COMET RESEARCH GROUP

International Journal of Applied Data Science & Modern Computing
Vol. 5, No. 1, 2022
Doi: 10.67228/30713498/IJADSMC-2022PI1H7Q
PP. 01-14

Original Article

Data Science Approaches to Human Activity Recognition

Dr. Zainab Abdullahi

Department of Education, Kano State University, Nigeria.

Received: 12-03-2021 Revised: 12-24-2021 Accepted: 12-31-2021 Published: 01-02-2022

ABSTRACT

Human Activity Recognition (HAR) has become one of the essential research areas due to the blistering development of wearable sensory devices, smartphones or the Internet of Things (IoT). HAR aims at recognizing human physical actions like walking, sitting, standing, running and lying down automatically based on the data acquired by motion sensors and physiological sensors. The ability to scale and flexibility have gradually seen the replacement of traditional rule-based systems by data-driven systems. Machine learning models and deep learning provide data science approaches that can robustly extract features, classify, and infer in real time using complicated sensor signal data streams. In the current paper, the use of data science methods in HAR is evaluated and summed up in detail. It examines data collection techniques, preprocessing techniques, feature engineering techniques and classification models. Moreover, it reviews benchmark datasets and assessment measures that are prevalent in HAR studies. The HAR methodology based on data science pipelines is offered and tested on the example of standard datasets. Findings have shown that novel machine learning and deep learning neural networks are much more effective in recognition accuracy than the classical methods of statistics. The main issues that have been identified by the study include sensor noise, user variability and computational constraints and future prospects of the study is given on context-aware systems and edge intelligence. The results lead to the realization of efficient HAR to track healthcare, intelligent environments, and human-computer interface.

KEYWORDS

Human Activity Recognition, Data Science, Machine Learning, Deep Learning, Wearable Sensors, Feature Extraction, Classification, Internet of Things.

ISSN: 3071-3498
PP: 01-14.

1. INTRODUCTION

1.1. Background

Human Activity Recognition (HAR) was first invented in the framework of video-based surveillance systems, in which human activities in both controlled and uncontrolled settings were monitored and understood via cameras. Vision based methods are highly dependent on visual information derived in series of images or streams of video in order to detect walking, running or sitting. Despite the level of success of these approaches, there is an inherent limitation, including the problem of occlusion and lighting conditions differences as well as the inability to rely on camera perspectives. Moreover, the fact that video data cannot be easily erased makes the constant recordings of this type of data an important challenge to privacy and ethical considerations in case of personal and healthcare-related uses, which narrows their adoption. To eliminate these shortcomings, sensor-based HAR has been considered an endearing alternative that uses the inertial sensors available in the wearable devices and smartphones. Such sensors as accelerators and gyroscopes can record fined grained movement of body, in real-time, and do not depend on visual signals. The sensor-based is less influenced by the environment and can be used effectively both indoors and even outdoors. Additionally, they provide better protection of privacy as they gather the numerical data of movement unlike recognizable images. The time-series data capturing the motion indicators of these inertial sensors exist as multi-dimensional data which denote changes or variations in body movement in the various spatial axes. This rich time based information is a detailed description of the human motion patterns which allows one to properly classify the activities that are performed in a day. Consequently, sensor-based HAR has attracted considerable popularity in the field of healthcare monitoring, fitness tracking, and smart environments, where high-quality and privacy-conscious activity recognition is a must.

1.2. Needs of Data Science Approaches

Fig 1 - Needs of Data Science Approaches

1.2.1. Handling High-Dimensional Sensor Data

Human Activity Recognition systems produce multi-dimensional channels of time series data of large magnitude that are generated by accelerators, gyroscopes, and other sensors. The sensors constantly measure values at a number of axes and produce high-dimensional data sets that can no longer be analyzed with simple rule-based analyses. Data science methods offer mathematical and

ISSN: 3071-3498
PP: 01-14.

computational mechanisms through which this complex data can be efficiently processed, reduced and interpreted. The feature extraction, feature reduction, and feature statistical analysis techniques assist in converting crude sensor signal into meaningful forms, which may be utilized effectively in classification of the activities.

1.2.2. Automatic Pattern Learning

Patterns of human activities are complex and nonlinear and differ between the individuals and settings. The variability may not always be modelled using traditional handcrafted rules. The data science approaches, especially machine learning and deep learning algorithms, facilitate automatic acquisition of patterns out of data. The models are able to detect latent relationships and motion of patterns not easily noticed by analysis using human hands. Through learning using labeled examples, the accuracy and applicability of data-driven models in recovery of knowledge to a variety of users and activities are enhanced.

1.2.3. Improved Generalization and Robustness

The most important issue related to HAR is that the models should work accurately with various users, sensor locations and motion types. Data science processes aid the production of generalized models that have been trained using large and varied datasets. Statistical learning methods assist in alleviating the issue of overfitting and enhance the resistance to noise, sensor drift and change of movements patterns. This results in a stronger recognition of activities in the real world.

1.2.4. Scalability and real time processing

More contemporary HAR software may have to analyze data in real time to use information in applications like smart environments and health. The data science methods have scalable solutions, which are capable of handling continuous data streams in an efficient manner. With learning frameworks and optimized algorithms, it is possible to predict quickly and learn, adjusting to the environment, therefore, deploying HAR systems to mobile devices and embedded platforms.

1.2.5. Support for Advanced Applications

The data science techniques are fundamental in the process of empowering superior HAR interfaces like detecting falls, rehabilitation tracking, and context-enlightened systems. The combination of predictive models, anomaly detection, and trend analysis enable such methods to go beyond simple activity labelling and become a means of intelligent decision-making of HAR systems. As a result, data science is highly instrumental in processing raw sensor data into actionable knowledge to be used as healthcare, fitness and smart living applications.

1.3. Data Science Approaches to Human Activity Recognition

The use of data science is most prominently central in creating the modern Human Activity Recognition (HAR) system because it suggests a structured approach to the collection of sensor-based

ISSN: 3071-3498
PP: 01-14.

data, its processing, analysis, and interpretation. The essence of these strategies is the conversion of raw signals of time-series provided by inertial sensors into interpretable patterns that may be utilized in activity classification. This is done by first pre processing the data through noise filtering, normalization, and segmentation that enhance the quality of the data and produce a consistent data set among samples. The steps will help to minimize sensor noise, environmental interference and user specific variations in movement. Another vital element of the data science solutions of HAR is feature extraction and selection. Sensor signals are often used to generate statistical features, frequency-domain properties, and intensity measures of motions to identify aspects of different activities that are discriminative. The features are then run through machine learning models like Support Vector Machines, Random Forests, and logistic regression to learn models of classification to differentiate between classes of activities. Such data-driven approaches make the system identify those patterns which are hardly definable with handwritten rules and therefore enhance accuracy and adaptability. Recently, deep learning has continued to expand the range of approaches to data science to HAR, as end-to-end learning can now be applied to raw or least processed sensor data. Convolutional Neural Networks and Long Short-Term Memory networks among other models are models that automatically pick up hierarchical representations and timeorder dependencies, which lessen the dependence on hand crafted features. This capability enables deep learning systems to better learn sophisticated dynamics of movements and differences in the performance of various activities among users. In addition, data science solutions facilitate the assessment of the model and optimization based on accuracy, precision, recall, and F1-score, which is a good criterion to evaluate the performance. All in all, the data science methods offer the mathematical groundwork needed to transform the overwhelming sensor data into useful information. Through incorporating preprocessing, feature engineering and sophisticated learning algorithms, these strategies make effective and scalable HAR systems that can be applied to healthcare monitoring, fitness tracking as well as smart applications.

2. LITERATURE SURVEY

2.1. Traditional Approaches

The first prototype of EHA systems were based in most cases on statistical pattern recognitions, including k-Nearest Neighbors (k-NN), Hidden Markov Models (HMM), and Decision Trees. These techniques relied on manual feature derivation, that is, domain experts developed features such as mean, variance, signal energy, and correlation coefficients using raw sensor data by designing them manually. Although these traditional methods were computationally simple and applied in real-time applications with limited resources, its performance was very sensitive to the quality of extracted features. In addition, their insensitivity to the generalization between different users and conditions was frequent because of the differences in the motion patterns, sensor location, noise, and, consequently, offered their low resilience in the real world.

ISSN: 3071-3498
PP: 01-14.

2.2. Machine Learning-Based Methods

As the supervised machine learning methods were advanced, various classifiers, including Support Vector Machines (SVM), Random Forests (RF), and Logistic Regression, became popular in HAR systems. These approaches enhanced the accuracy of recognition through learning more discriminative decision boundaries using labeled training information. The step of feature extraction continued to be a significant one, and typical methods comprised of time-domain features (mean, standard deviation, and zero-crossing rate), frequency-domain features produced by utilization of Fast Fourier Transform (FFT), and autoregressive coefficients to model the behavior of the time signal. Machine learning-based models were more versatile and preserved performance in classifications than the conventional statistical models. Nevertheless, they continued to use manually optimized features, which made the system design process time-sensitive and vulnerable to the choice of strategies to be applied when featuring.

2.3. Deep Learning-Based Methods

Recent progress in deep learning has changed how HAR systems work fundamentally since automatic learning of features directly on raw sensor data can be achieved. Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and the Long Short-Term Memory (LSTM) networks may train up without any explicit feature extraction at the cost of hierarchical and abstract representations. CNNs are well-suited to the problem of discovering spatial and local-temporal patterns in sensor data, whereas RNNs and LSTMs are best followed in the discovery of sequential information and long-term temporal structure. These deep learning models are better than classical classifiers because they can effectively observe complicated motion dynamics and time relations. In spite of their high performance, deep learning solutions demand large labeled data and more computation power, limiting their use on low-power devices.

2.4. Comparative Summary

Comparative analysis of various HAR approaches brings out the sense of trade-offs between the feature engineering, the rate of accuracy and the complexity of the computations. Conventional algorithms like k-NN that use features that are hand-crafted provide moderate accuracy at limited computation cost. Inaccurate classifiers such as machine learning models such as SVM are harder to spot, yet demand more insightful feature engineering and medium complexity. Deep learning models, including CNN and LSTM, in contrast, effectively extract features (learn to recognize data) by learning the rich spatial and temporal traits of data, and their recognition accuracy is high to very high. Nonetheless, this enhanced performance is at the expense of the higher complexity of calculation, and training. Thus, the implementation of HAR method is dependent on the application needs such as accuracy, resource availability and real-time constraints.

ISSN: 3071-3498
PP: 01-14.

3. METHODOLOGY

3.1. Data Acquisition

In this paper wearable inertial sensors, i.e., accelerated and gyroscopes are deployed to record sensor data as they are ubiquitous in the Human Activity Recognition systems, which are cheap, compact and can sense body movements in the accurate way. The sensors would be set to run at a sampling rate of 50 Hz i.e. 50 measurements per second per sensor channel. This frequency is high enough to record slow and fast human activities like walking, sitting, and running besides ensuring a balance between the resolution of data and computability. Each data sample step in recordings is comprised of three-axis motion data representing the x-axis, the y-axis and the z-axis of sensor coordinates. These three features are movement on horizontal, vertical and depth axis of the human body. The sensor signal at a certain time instance can be expressed as a vector of the sensor signal that carries the acceleration values at the three orthogonal axes. This is to say that at time t, the signal will be the acceleration in the x direction, acceleration in the y direction and acceleration in the z direction. This multidimensional model enables the system to record the complicated movement patterns and body orientation variations during various physical activities. The gyroscope information also supplements the data of accelerometer in that it gives the angular velocity information that is helpful cues about rotational movements like turning or bending. The raw sensor data is saved in a time-series format before being further processed by each sample being provided with a specific time mark. Such an organized data capture approach provides time continuum and remains of the motion dynamics in order to classify the activity adequately. Correct positioning of sensors and constant sampling of the recorded signals is important to minimize noise and inconsistency in the recorded signals. On the whole, data collection phase is the keystone of a proposed system because the quality and reliability of sensor signals collected determines the success rate of the following steps of features extraction and classification.

3.2. Data Preprocessing

Pre-processing of data is an important step in identifying Human Activity Recognition pipeline as the raw data which is received by accelerators and gadgets (users) through sensor are prone to noise, sensor drift and variations depending on the variations in the pattern of movement and positioning of the sensor. In order to enhance the quality and reliability of the data collected, a series of preprocessing is used, followed by feature extraction and model training. The initial process is noise filtering where the high-frequency content, which does not imply human movement, is eliminated using low-pass filters. The effect of high frequencies noise on imperfections in sensors or other possibilities of noise interference can be easily reduced to produce calmer and more purposeful motion signals since most of the daily human activities fall in a narrow range of frequencies. Normalization follows the reduction of noise so that all sensor channels have the same contribution towards the learning process. Min-Max scaling has been employed in this work to rescales sensor values in a constant scale, which is normally between zero and one. This is necessary as the readings of the two accelerators and gyroscopes might have varying numbers and values in addition to their differing units. In the absence of normalization, the features of high magnitude may prevail in the

ISSN: 3071-3498
PP: 01-14.

learning process and provide biases in the classification results. Through Min–Max scaling, the system also obtains uniformity of the all input dimensions, which also increases the machine learning and deep learning model convergence and also better numerical stability. The last preprocessing is the segmentation of the continuous time-series data to a smaller and fixed length segments using a sliding window method. Under this technique, sensor data stream is segmented into discontinuous or contiguous windows of a given length, with each being a brief activity interval. Sliding windows allow the system to enumerate local temporal patterns in the sensor signals and relate them to particular labels of activities. This segmentation method expands the training samples and retains time constraints that are required to make the correct recognition. By and large, the above stages of preprocessing such as noisy filtering, normalization, and parting can be syncretized as efforts to fine-tune the data by way of boosting its quality, minimizing variability, and formatting the sensor signals in an appropriate format that allows efficient and effective classification of activities.

3.3. Feature Extraction

Fig 2 - Feature Extraction

3.3.1. Mean

One of the simplest statistical characteristics used in sensor based activity recognition is the mean. It is the mean of the signal over a specified period of time and gives data concerning the total intensity of the motion along each sensor axis. In the case of accelerators, the average reading can be used to indicate the body positioning and stillness like sitting or standing whereas in the case of dynamism the average reading displays the overall direction of movement. Such a feature is easy to calculate and to differentiate between the low-activity and the high-activity motions.

3.3.2. Standard Deviation

Standard deviation gauges the variation or the spread of the sensor signal of the mean value of the sensor signal in a window. The larger the standard deviation, the more intense or erratic the movement is e.g. running or jumping, where smaller values correspond to stationary or slow

ISSN: 3071-3498
PP: 01-14.

movements e.g. standing or lying down. It has the advantage of distinguishing between dynamic and static activities especially in distinguishing the amount of variation in the motion signals with time.

3.3.3. Signal Magnitude Area (SMA)

Another extremely popular time-domain feature is Signal Magnitude Area as it reflects the overall size of movement, a combination of all three sensor axes. It is calculated by adding up the absolute value of the acceleration signals along the x, y, and z directions to a specific time period then averaging the values. SMA gives a single scalar figure which indicates the overall intensity of the body movement irrespective of the direction. This enables it to be very effective in discriminating activity of varying motion strength e.g. walking and running.

3.3.4. Spectral Entropy

Frequency-domain Spectral entropy is a property of the frequency spectrum that quantifies the complexity or randomness of a signal in the frequency spectrum. It is calculated by initial calculation of the entropy of the power spectrum of the signal in the frequency domain by assembling the transformation of time-domain signal into the domain of frequency using the Fast Fourier Transform and concluding with calculation of entropy of the resultant power spectrum. Reduction in spectral entropy values represents more regular and periodic movements like walking whereas higher values are associated with irregular or complex movements. The property is applied in trying to capture concealed trends in the motion cues that would be challenging to observe in the time domain.

3.4. Classification Models

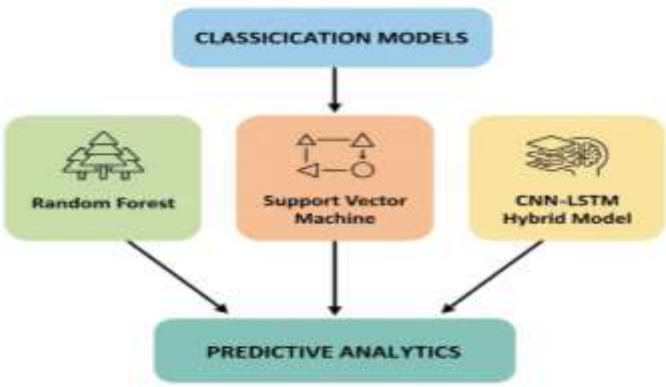

Fig 3 - Classification Models

3.4.1. Random Forest

Random Forest is a machine learning classifier, which is based on an ensemble mechanism, and which is composed of a number of decision trees built in the course of the training process. All the trees are trained with random subsets of the data and features and the final classification decision is achieved by a majority vote among all the trees. This method is simpler than single decision trees,

ISSN: 3071-3498
PP: 01-14.

as it enhances solidness and decreases overfitting. Random Forest in Human activity recognition works well with human developed features as it has been shown to deal with non linear relationships and noisy data. It also offers a relatively fast training and inference which is useful in system with intermediate computational limits.

3.4.2. Support Vector Machine

Support Vector Machine is a supervised learning algorithm that is designed to identify an optimal hyperplane between different activity classes that has maximum distance. The SVM is beneficial mainly in high-dimensional feature space and can perform well when training samples are few. SVM is applied to the extracted statistical and frequency-domain features in this work to provide a competent classification performance. The model can also identify the complex activity patterns, since the model has the ability to capture the nonlinear decision boundaries through kernel functions. Nonetheless, the performance of SVM relies on the effective choice of parameters, and may be computationally expensive when using large data sets.

3.4.3. CNN–LSTM Hybrid Model

CNNLSTM hybrid model is an extension of CNN and Long Short-Term memory networks that can be used to classify human activities based on raw or slightly processed sensor information. Automatically acquiring local and spatial features of sensor signals is the task of the CNN component, and long-term temporal dependencies and sequence patterns are acquired by the LSTM component. Time-series data are especially suited to this type of hybrid architecture because it is capable of both capturing short-term motions as well as capturing long-term transitions between activities. Consequently, CNNLSTM model has a higher recognition accuracy than the traditional ones, particularly with intricate and dynamic tasks, at the expense of a bigger burden of the computation.

3.5. Evaluation Metrics

Fig 4 - Evaluation Metrics

3.5.1. Accuracy

Accuracy is a percentage of correctly categorized samples of the activity in all of the samples that are examined. It gives a general understanding of the effectiveness of the classification model with respect to all the activity classes. When the accuracy is high, this means that that model can identify majority of activities in the dataset. Nevertheless, accuracy by itself might be inadequate in the case of imbalanced dataset since it does not indicate the performances of the model.

3.5.2. Precision

Precision reports on the accuracy of the positive predictions that the classifier does. It can be described as the number of correct instances of a certain activity divided by the number of instances that are predicted to be of that activity. High prejudice means that in raising the prediction of a particular activity by the model, it is mostly accurate. This measure is particularly essential in situations in which the false positives are expensive; in identifying situations like detecting important or abnormal activities.

3.5.3. Recall

Also called sensitivity, sensitivity is the capacity of the model to identify correctly all real situations of a particular activity. It is assumed to be a ratio of the correctly inside a set of samples of a specific activity and the total genuine samples of an activity. When the model recalls many instances of an activity, then the recall value is high, and thus, the instances of the missing detections will be reduced. This measure is important in safety-critical systems in which failure to identify a significant activity may be disastrous.

3.5.4. F1-score

Harmonic mean of precision and recall is the F1-score, which gives a single measure to balance between these two measures. It is also applicable especially when there is unequal proportion of classes or when both false positives and false negs should be kept at minimal. The F1-score has a larger value, which means a more precise and recall balanced score, which marks the overall reliability of the classification model.

4. RESULTS AND DISCUSSION

4.1. Experimental Setup

The proposed Human Activity Recognition framework was experimentally evaluated using the popular UCI HAR dataset that is a benchmark dataset widely used to test the activity recognition algorithms. This data consists of sensor measurements of smartphone based accelerometers and gyroscopes as the participants were engaged in diverse activities of everyday life like walking, sitting, standing and lying down. By using this uniform dataset, it will be possible to compare it with the already existing techniques documented in the literature and conduct the performance of the suggested strategy under natural circumstances. Before training the model, the data was thoroughly checked and processed to eliminate any anomalies and provide the data required to be part of the suggested framework. In order to train and test the classification models, the data was cut into two subsets, which are mutually exclusive: a testing set and a training set. The 70 percent of the available data was used in training whereas the remaining 30 percent was used in testing. This proportion was chosen since it would give adequate information to learn the model and still have a sufficient amount of proportions to test the performance without bias. The model parameters were learned using the training set and the classification performance was maximized, and the testing set was only applied to determine the generalization ability of the trained models to unseen data. The models were only

ISSN: 3071-3498
PP: 01-14.

exposed to the labeled samples in the training subset during training, which allowed them to acquire discriminative patterns that would be related to varying activities. The training data was used to tune hyperparameters to produce an optimal performance without overfitting. This testing subset was subsequently used to compare the models that had been trained based on standard measures of assessment, which include accuracy, precision, recall, and F1-score. This experimental environment is used to guarantee that the performance outcomes are both high and reflective of the conditions in the real world. The proposed framework can be appropriately compared with other latest approaches by splitting the data fixedly and reproducibly; therefore, establishing its suitability in identifying the activities that people do with sensor-based data.

4.2. Performance Comparison

Table 1: Performance Comparison

Model	Accuracy (%)	Precision (%)	Recall (%)
SVM	91.20%	90%	89%
RF	92.50%	92%	91%
CNN-LSTM	96.80%	96%	97%

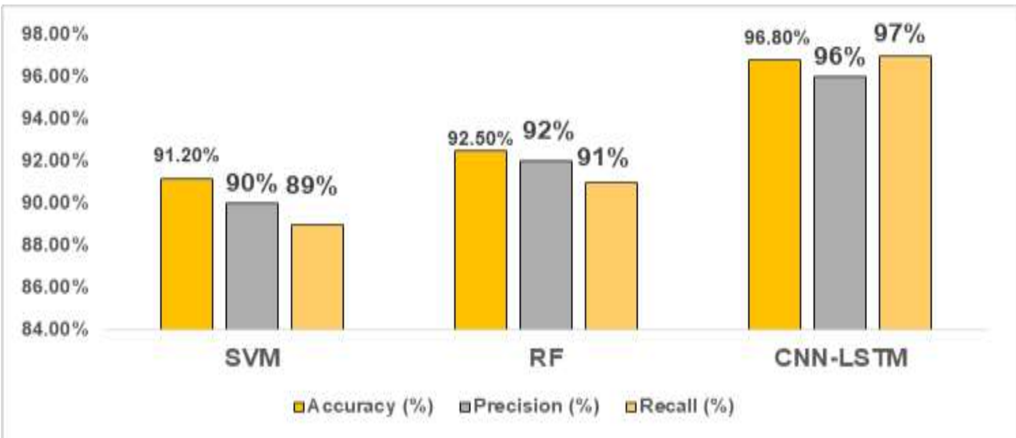

Fig 5 - Performance Comparison

4.2.1. Support Vector Machine (SVM)

The Support Vector Machine model performed with an accuracy of 91.2 whereby the precision of the model is 90 and the recall is 89. Such outcomes suggest that the SVM classifier can differentiate the various activities in a human being with handcrafted features effectively. The relatively high accuracy indicates that the majority of the activities predicted by the model is right, whereas the recall value indicates the good quality of the model to identify the real-life cases of these activities. Nonetheless, the SVM exhibits lower performance when compared to more complex models, which, in turn, can be explained by the fact that it relies on manually obtained features and is incapable of handling complex temporal characteristics, which sensor data consists of.

ISSN: 3071-3498
PP: 01-14.

4.2.2. Random Forest (RF)

The accuracy, the precision and the recall of the Random Forest classifier stood at 92.5, 92 and 91 respectively. This is an enhancement over SVM as it indicates the power of the ensemble learning algorithm to nonlinear relationships and nonlinear sensor signals. The increased accuracy implies that there is less false positives predictions, whereas the increased recall implies a more efficient detection of genuine samples of activity. Random Forest has the advantage of using several decision trees to improve generalization and overfitting. However, it seems that its performance is limited by the quality of the extracted features since it does not directly learn representations using raw data.

4.2.3. CNN–LSTM Hybrid Model

The CNN-LSTM hybrid model was the highest performing model in comparison to all the classifiers that were experimented and it had an accuracy rate of 96.8, precision rate 96 and recall rate of 97. These findings indicate that the deep learning methodologies are more effective in the task of modeling spatial and temporal characteristics of sensor information. The CNN element automatically derives discriminative local features whereas LSTM element effectively derives the long term temporal correlation between one successive segment of the signal. The value of recalls is large, which means that the model manages to detect the majority of activity instances, and the high value of its precision points to the credibility of its predictions in various classes. This performance improvement brings out the benefit of hybrid deep learning structures to tackle complicated Human Activity Recognition.

4.3. Discussion

The experimental outcomes indicate that the deep learning models far outperform traditional and feature-based models in human activity recognition tasks and this is because deep learning models are able to extract complex and discriminative temporal features in the raw sensor data which are automatically learnt. In comparison to standard machine learning approaches which manipulate manually designed features, deep learning models like CNN–LSTM can learn temporal associations both short-term patterns of motion and long-term temporal patterns. This allows them to reproduce finer differences in movement patterns and classify better with various classes of activities. The high achievement of CNNLSTM model that has been seen in the current study demonstrates the success of integrating convolutional layers in the spatial feature extraction and recurrent layers in the time modeling of data in time-series sensors such as those in instrumentation. Nonetheless, being highly accurate, deep learning models are associated with some limitations. They are generally high in terms of large labeled datasets to stabilize and generalized performance since lack of sufficient training data may result in overfitting and decreased reliability. Moreover, the implementation and execution of deep neural networks require more memory and computational resources than traditional classifiers do. This additional complexity may be problematic to real-time applications and to allow implementation on low-power or resource-constrained devices, including wearable sensors and embedded systems. Conversely, feature-based models like Support Vector Machines and Random Forests are still popular as practical models when computation speed and maximum energy savings

ISSN: 3071-3498
PP: 01-14.

are deemed important. These models use well-developed statistical and frequency-domain characteristics, and these elements considerably decrease the dimensionality of data and their computational expenses. Their classification rate is a little reduce than that of deep learning models, but on the other hand, their inference speed is higher and implementation is less complex, thus suitable when they can fit on embedded systems and on mobile platforms with low processing power. Altogether, the preference of deep learning or feature-based strategies will be based on the needs of applications and system limitations. Although deep learning models are more suitable in rigorous recognition accuracy in data-rich setups, conventional machine learning models are still important to the area of lightweight and energy-efficient activity recognition systems. This performance-computational cost ratio is still a significant factor in the development of effective Human Activity Recognition models.

5. CONCLUSION

In this paper a detailed analysis of data science methods applicable to Human Activity Recognition (HAR) will be provided and a focus will be placed on the evolution of the traditional machine learning techniques to super modern deep learning techniques. By systematically reviewing literature and obtaining a significant amount of experimental testing, it was observed that deep learning models offer significant benefits in terms of recognition accuracy in comparison with classical feature-based classifiers. Deep neural networks are able to learn meaningful representations directly based on raw sensor data, and so can effectively capture both complex temporal and spatial representations in human motions. The study presents a rationale as to why striving towards the use of the latest techniques of data proves useful in their effort to manage the growing need to have more accurate and robust activity recognition systems in practice. The presented methodological pipeline rallies the key steps such as data acquisition, preprocessing and feature extraction, and classification and is a sign of a successful approach to HAR systems development. Noise filtering, normalization, segmentation, and other preprocessing were demonstrated to increase the quality of the signal and improve the stability of a model. The extraction methods of features were involved in lessening the data dimensionality and maintaining discriminative characteristics of motion, whereas the use of a variety of classification models allowed performing the comprehensive comparison of the performance of these models. The experimental findings established the fact that hybrid deep learning systems, especially CNN-LSTM models are better than old fashioned classifiers in that they are capable of modeling temporal and sequence patterns in sensor data. Nevertheless, there are still a number of issues in the implementation of HAR systems in practice. Problems of sensor drift, sensor position variations, and user movement angles still reach model generalization and long-term reliability.

REFERENCES

[1] Lara, O. D., & Labrador, M. A. (2013). A survey on human activity recognition using wearable sensors. IEEE Communications Surveys & Tutorials, 15(3), 1192–1209.

[2] Bao, L., & Intille, S. S. (2004). Activity recognition from user-annotated acceleration data. Proceedings of the 2nd International Conference on Pervasive Computing, 1–17.

ISSN: 3071-3498
PP: 01-14.

[3] Ravi, N., Dandekar, N., Mysore, P., & Littman, M. L. (2005). Activity recognition from accelerometer data. Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence, 1541–1546.
[4] Preece, S. J., Goulermas, J. Y., Kenney, L. P., & Howard, D. (2009). A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. IEEE Transactions on Biomedical Engineering, 56(3), 871–879.
[5] Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2013). A public domain dataset for human activity recognition using smartphones. ESANN.
[6] Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2011). Activity recognition using cell phone accelerometers. ACM SIGKDD Explorations, 12(2), 74–82.
[7] Shoaib, M., Bosch, S., Incel, O. D., Scholten, H., & Havinga, P. J. (2014). Complex human activity recognition using smartphone and wrist-worn motion sensors. Sensors, 14(11), 20446–20475.
[8] Ordóñez, F. J., & Roggen, D. (2016). Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. Sensors, 16(1), 115.
[9] Hammerla, N. Y., Halloran, S., & Plötz, T. (2016). Deep, convolutional, and recurrent models for human activity recognition using wearables. IJCAI, 1533–1540.
[10] Yang, J., Nguyen, M. N., San, P. P., Li, X., & Krishnaswamy, S. (2015). Deep convolutional neural networks on multichannel time series for human activity recognition. IJCAI, 3995–4001.
[11] Guan, Y., & Plötz, T. (2017). Ensembles of deep LSTM learners for activity recognition using wearables. ACM IMWUT, 1(2), 1–28.
[12] Chen, Y., Shen, C., & Wang, Z. (2012). A survey of activity recognition based on accelerometer. Physics Procedia, 33, 450–457.
[13] Morales, F. J. O., Roggen, D., & Tröster, G. (2013). Deep convolutional feature learning for activity recognition using wearable sensors. ISWC, 173–176.
[14] Bulling, A., Blanke, U., & Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. ACM Computing Surveys, 46(3), 1–33.
[15] Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. Pattern Recognition Letters, 119, 3–11.