

A Survey of Machine Learning Techniques For Human Activity Recognition and Their Methods and Algorithm

R. Bagavathi Lakshmi¹ and S. Parthasarathy^{2*}

¹*Research and Development Centre, Bharathiar University, Coimbatore, India*

²*Department of Computer Applications, Thiagarajar College of Engineering, Madurai, India*

*Corresponding author: parthatce@gmail.com

ABSTRACT

Human Activity Recognition (HAR) refers to the automatic detection of various physical events carried out by people in their daily life. HAR system assists in differentiating the behavior performed by person and informative feedback is offered for involvement. The main objective of activity recognition is to offer information on a user's actions for permitting the computing systems to proactively help users with their tasks. The recognition model is constructed and concentrated on different machine learning algorithm to set the feature illustration in this research work based upon the efficient of HAR. But, HAR failed to reduce the error rate during human actions identification. In order to enhance the performance of HAR, the accuracy in identifying human activities needs to be raised.

Keywords: Human Activity Recognition, detection of human actions, discriminative features, machine learning techniques.

HAR is an automatic discovery and detection of actions from the investigation of data obtained by sensors. The learning of individual identity, behavior and activities developed by one or more persons in video sequences serves significant for diverse applications. HAR discover probable applications in number of sectors, beginning from surveillance responsibilities for security to assistive living and developing the quality of life or gaming. In addition, HAR helps in progressing and spreading of dominant and reasonable mobile devices or wearable sensors.

The extraction of action data for predicting and characterizing the actions is one of the significant tasks in activity recognition is used in diverse sectors like personal healthcare. Machine learning algorithms are artificial intelligence techniques which help in characterizing the data in an efficient manner to extract the knowledge and activity recognition system to make intelligent decisions. In HAR, selection of appropriate

informative data is a demanding task. The error in informative data leads to dissemination of error to the model which recognizes the activity. In addition to, HAR also involves high computational necessities. An efficient detection of discriminative features is essential in recognizing the human activities.

This paper is structured as follows: Section II reviews on existing machine learning techniques used in identifying human activities, Section III gives brief discussion about machine learning techniques for human activity recognition, Section IV identifies the possible comparison between them, Section V explains the limitations as well as the related work and Section VI concludes the paper, key areas of research is given to enhance the human activity recognition accuracy with minimum computational complexity and error rate in extracting the discriminative features.

LITERATURE REVIEW

A robust human activity recognition system was developed in^[1] depended on coordinate transformation and principal component analysis (CT-PCA) and online support vector machine (OSVM). This assisted in eliminating the consequence of orientation variations. An online independent support vector machine (OISVM) was developed to increase the activity recognition accuracy. But, feature extraction was not considered for enhancing the performance of activity recognition.

An intuitive method was introduced in^[2] to preserve the geometric construction of significant action snippets within skeleton sequence. Encoding of temporal information was performed by employing a dividing algorithm. Posture tendency descriptor (PTD) was implemented for characterizing one action snippet. At last, various PTDs with whole skeleton sequence were combined in a hierarchical and temporal order to produce the illustration of a human activity. However, the detection of abnormal activities was not performed.

A user-independent deep learning-based approach was implemented in^[3] for online human activity categorization. Convolutional Neural Networks were developed to mine local features with simple statistical features to conserve data about global form of time series. The running time was decreased and capable of implementation on mobile devices in real time. But, computational performance was more complex. A deep learning model was discussed in^[4] for categorizing human activities without the utilization of preceding information. However, the error rate during detection was not reduced efficiently.

A low-cost, unobtrusive and robust system was developed in^[5] to maintain self-govern existing of elder people. The actions of a person were recognized through deciphering signal fluctuations with the aid of radio-frequency identification (RFID) and machine learning approaches. Then, a compressive sensing, dictionary-based technique was employed for handling noisy, streaming and uneven RFID signals for recognizing informative dictionaries of actions by means of unsupervised subspace disintegration. The characteristics of sparse coefficients were developed to employ the embodied discriminative data on action detection task. Efficient activity detection was attained through the illustration of actions. The performance of the system was not improved for dynamic conditions.

MACHINE LEARNING TECHNIQUES FOR HUMAN ACTIVITY RECOGNITION

Human Activity Recognition (HAR) is the supervising of people for identifying their physical behavior

to facilitate them in daily job or prevent emergencies. HAR system is developed to follow particular tasks to recover activity information from sensor. Activity Recognition Process comprises data acquisition, preprocessing, segmentation, feature mining and categorization. Identification of body activities is based on features extracted from motion investigation. The detection of discriminative features in human activity recognition is a demanding task because of its complexity. Machine learning is a relevance of artificial intelligence (AI) which helps the systems to perform repeated recognition and enhances from knowledge without being planned. Some of the recent related works regarding the human activity recognition are reviewed.

Principal Component Analysis and Online Support Vector Machine for Human Activity Recognition

Human activity recognition is one of the leading concepts employed in wearable devices or smartphones for supporting diverse applications namely, healthcare, smart devices etc. The utilization of smartphone incorporated with number of sensors causes degradation of recognition accuracy with orientation, assignment and subject difference. A robust human activity recognition system was introduced to resolve the problems associated to human activity recognition. The introduced recognition system was depended on coordinate transformation and principal component analysis (CT-PCA) and online support vector machine (OSVM). CT-PCA scheme was employed in restricting the dissimilarity due to orientation. Subsequently, online independent support vector machine (OISVM) was designed to increase the recognition accuracy. OSVM algorithm also used even a tiny fraction of data from subject for updating the parameters of SVM algorithm.

A 3D linear acceleration was utilized in modern smartphones. The orientations of smartphone based system vary with clothing, body shape and shifting rotation of device during dynamic actions. The outputs of smartphone sensors depend on device coordinate system and the data of three dimensions changes with varying device direction. This in turn corrupts the recognition accuracy of activity recognition scheme. Therefore, the sensor outputs were relocated from device coordinate system into a fixed coordinate system. The coordinate transformation approaches were depended on Euler angle and quaternion. Quaternion which was a four-element vector was employed in CT-PCA for characterizing any alternation in a 3D coordinate system.

After CT-PCA, 3D acceleration time series data is categorized into windows during feature mining. Then, time and frequency domain features were extracted from these windows. Followed by it, classification was performed to distinguish the discrete category of an observation. With the aid of OISVM, typical SVM outcomes were converged by including new observations with the aid of non-sensitive user-defined parameter. This helped in minimizing time and space requirements with higher accuracy.

Real-time Recognition of Human Activities Using Convolutional Neural Networks

In healthcare monitoring and fitness tracking, human activity recognition is used for obtaining instant activity recognition of a user. A user-independent deep learning-based approach was developed for online human activity categorization. Convolutional Neural Networks (CNN) is designed to mine the local features using statistical features for protecting data about global form of time series. Then, the effects of time series length on recognition accuracy in continuous real-time activity categorization was examined. Quasi-periodic accelerometer time series were considered by using CNN which applied filters

to small portions of data for extracting local data patterns and difference. Unsupervised feature learning was executed in the convolutional layers and the acquired features were passed to fully-linked layers for performing categorization. The convolutional layers were optimized with a typical back propagation method to reduce the prediction error. In addition to, computations and running time was also decreased by using CNNs.

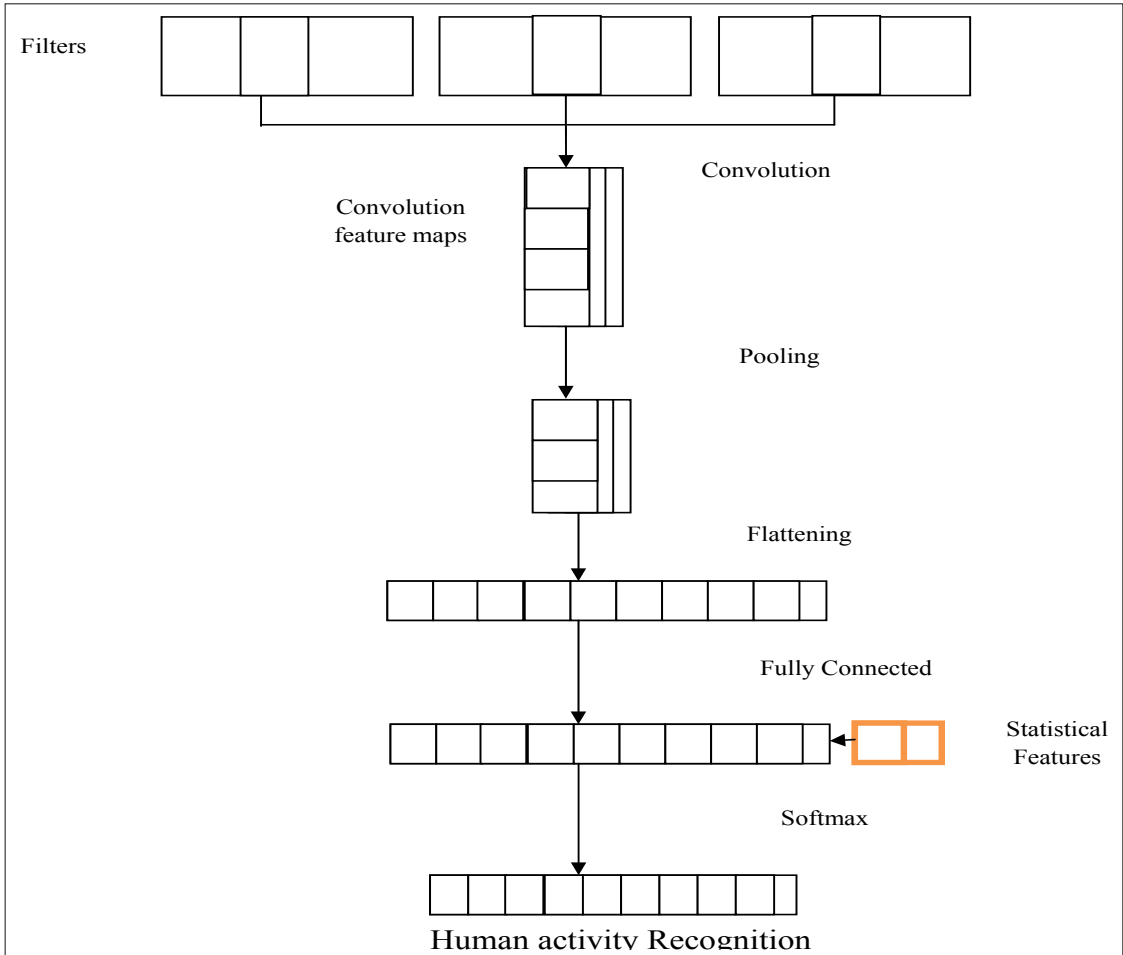


Fig. 1: Architecture of Convolutional Neural Networks for local feature extraction

As shown in Fig. 1, the centered accelerometer data was processed in the convolutional layer with convolutional filters for generating a rich feature characterization of data. Then, ReLU function was employed to the resulting feature maps and feature representation was decreased by using max-pooling. The output of max-pooling layer was compressed and loaded up together with supplementary statistical characteristics. Then, joint vector was passed to a fully-linked layer. Followed by it, dropout approach was used for restricting overfitting. At last, outputs of fully-linked layer were passed to a soft-max layer to evaluate probability distribution. Cross-entropy loss function was decreased with minimum computational cost.

Recurrent Neural Network in Recognition of Human Activities

In smart circumstances, human activity recognition with smart home sensors is one of the fundamental concepts in ubiquitous computing. A deep learning model was designed to categorize human actions. Then, Long Short Term Memory (LSTM) Recurrent Neural Network was employed in recognizing the human actions such as cooking, bathing and sleeping. LSTM is a recurrent neural network framework developed for handling gradient decay. LSTM layer's key element was memory block comprised with three gates such as input, output and forget gates provided for write, read and reset functions. Then, the information between each LSTM block was carried with the aid of an LSTM cell state and the variation of this cell state were maintained using three gates.

SCOMPARISON OF DIFFERENT MACHINE LEARNING TECHNIQUES FOR HUMAN ACTIVITY RECOGNITION & SUGGESTIONS

In order to compare machine learning methods for human activity recognition, number of activities and sequences are taken to perform the experiment. Various parameters are used for machine learning methods for feature extraction in enhancing the performance of human activity recognition.

Error Rate

Error rate is defined as the ratio of number of activities incorrectly identified to the total number of activity and is measured in terms of Percentage (%). It is mathematically formulated as given below:

$$ER = \frac{\text{No.of activities incorrectly identified}}{\text{Total no.of activity}} * 100$$

When error rate is lower, the method is said to be more efficient.

Table 1: Tabulation for Error Rate

Number of activities	Error Rate (%)		
	CT-PCA	CNN architecture	LSTM Recurrent Neural Network
10	15	18	21
20	18	22	24
30	20	24	26
40	22	25	29
50	25	27	30
60	27	30	32
70	28	33	34
80	30	35	38
90	31	37	39
100	32	40	42

Table 1 shows the error rate with respect to number of activities ranging from 10 to 100. Error rate comparison takes place on existing Coordinate Transformation and Principal Component Analysis (CT-PCA), Convolutional Neural Networks (CNN) Architecture and Long Short Term Memory (LSTM) Recurrent Neural Network. From the table value, it is clear that the error rate using CT-PCA is lower when compared to CNN Architecture and LSTM Recurrent Neural Network. The graphical analysis of error rate is shown in Fig. 2.

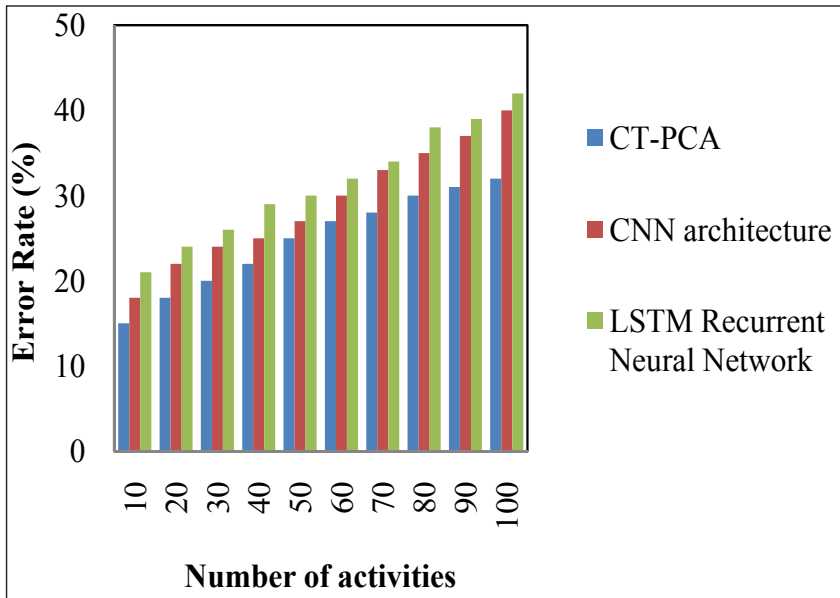


Fig. 2: Measure of Error Rate

As shown in Fig. 2, error rate during human activity recognition for different number of activities is evaluated. Coordinate Transformation and Principal Component Analysis (CT-PCA) has less error rate than that of Convolutional Neural Networks (CNN) Architecture and Long Short Term Memory (LSTM) Recurrent Neural Network. When the number of activities gets increased, attack detection rate gets decreased correspondingly in all three methods. But comparatively, CT-PCA reduces error rate by 15% than CNN Architecture and 22% than LSTM Recurrent Neural Network.

Computational Complexity

Computational complexity (CC) is defined as the product of number of sequences with the consumption of memory for one sequence. It is measured in terms of kilobytes (KB). It is mathematically formulated as given below:

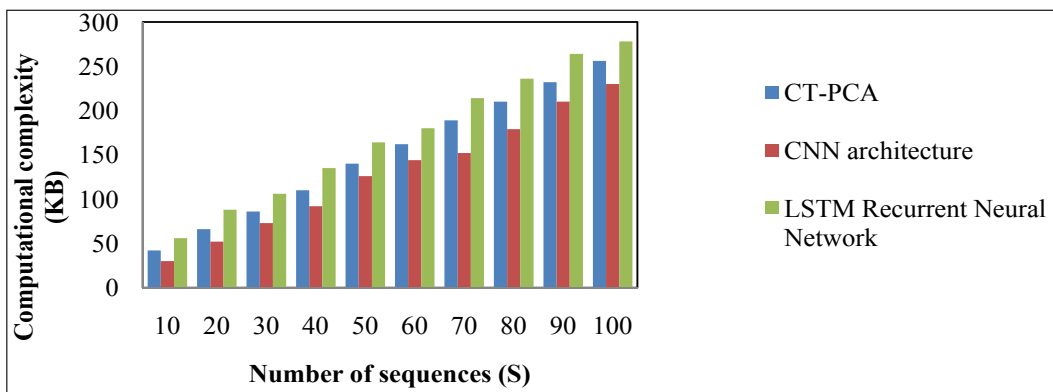
$$CC = \text{Number of sequences} * \text{memory consumption (one sequence)}$$

When computational complexity is lower, the method is said to be more efficient.

Table 2: Tabulation for Computational Complexity

Number of sequences (S)	Computational complexity (KB)		
	CT-PCA	CNN architecture	LSTM Recurrent Neural Network
10	42	30	56
20	66	52	88
30	86	73	106
40	110	92	135
50	140	126	164
60	162	144	180
70	189	152	214
80	210	179	236
90	232	210	264
100	256	230	278

Table 2 shows the computational complexity with respect to number of sequences ranging from 10 to 100. Computational complexity comparison takes place on existing Coordinate Transformation and Principal Component Analysis (CT-PCA), Convolutional Neural Networks (CNN) Architecture and Long Short Term Memory (LSTM) Recurrent Neural Network. From the table value, it is clear that the computational complexity using CNN Architecture is lower when compared to CT-PCA and LSTM Recurrent Neural Network. The graphical analysis of computational complexity is shown in Fig. 3.

**Fig. 3:** Measure of Computational Complexity

As shown in Fig. 3, computational complexity during human activity recognition for different number of activities is evaluated. Convolutional Neural Networks (CNN) Architecture has less computational complexity than that of Coordinate Transformation and Principal Component Analysis (CT-PCA) and Long Short Term Memory (LSTM) Recurrent Neural Network. When the number of activities gets increased, computational complexity gets decreased correspondingly in all three methods. But considerably, CNN Architecture reduces computational complexity by 16% than CT-PCA and 28% than LSTM Recurrent Neural Network.

Human Activity Recognition Accuracy

Human Activity Recognition Accuracy is defined as the ratio of number of sequence successfully recognized to the total number of sequence. It is measured in terms of Percentage (%). It is mathematically formulated as given below:

$$HARA = \frac{\text{No.of sequence successfully recognized}}{\text{Total no.of sequence}} * 100$$

When human activity recognition accuracy is higher, the method is said to be more efficient.

Table 3: Tabulation for Human Recognition Accuracy

Number of sequences (S)	Human Activity Recognition Accuracy (%)		
	CT-PCA	CNN architecture	LSTM Recurrent Neural Network
10	50	60	70
20	45	65	75
30	63	77	87
40	58	75	88
50	64	82	92
60	65	83	92
70	73	89	97
80	74	88	96
90	73	83	91
100	82	90	92

Table 3 shows the human activity recognition accuracy with respect to number of sequences ranging from 10 to 100. Human activity recognition accuracy comparison takes place on existing Coordinate Transformation and Principal Component Analysis (CT-PCA), Convolutional Neural Networks (CNN) Architecture and Long Short Term Memory (LSTM) Recurrent Neural Network. From the table value, it is clear that the human activity recognition accuracy using LSTM is higher when compared to CT-PCA and CNN Architecture. The graphical analysis of computational complexity is shown in Fig. 4.

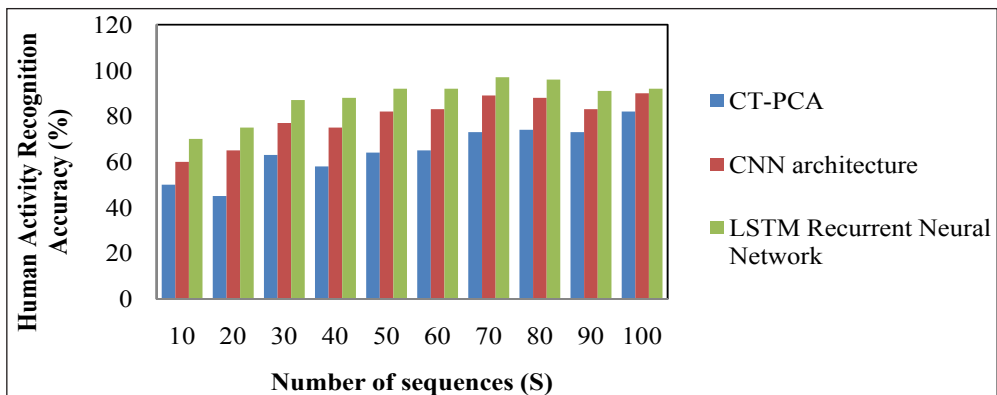


Fig. 4: Measure of Human Activity Recognition Accuracy

As shown in Fig. 4, human activity recognition accuracy during human activity recognition for different number of activities is evaluated. Long Short Term Memory (LSTM) Recurrent Neural Network has higher human activity recognition accuracy than that of Coordinate Transformation and Principal Component Analysis (CT-PCA) and Convolutional Neural Networks (CNN) Architecture. When the number of activities gets increased, human activity recognition accuracy also gets increased correspondingly in all three methods. Therefore, LSTM increases human activity recognition accuracy by 38% than CT-PCA and 12% than CNN Architecture.

DISCUSSION ON LIMITATION OF MACHINE LEARNING TECHNIQUES FOR HUMAN ACTIVITY RECOGNITION

A human activity recognition system was introduced using coordinate transformation and principal component analysis (CT-PCA) scheme for removing the impacts of orientation distinction. Online independent support vector machine (OISVM) algorithm was developed for attaining superior recognition performance. However, feature extraction and selection approaches were not performed.

A user-independent human activity recognition issue was addressed depended on Convolutional Neural Networks (CNN). Only minimum recognition intervals were required and implementation on mobile devices in real time was performed. But, CNN failed in considering the computational complexities. Long Short Term Memory (LSTM) Recurrent Neural Network was developed to categorize human actions but the variance during predictions was not minimized efficiently.

Related Works

A human activity recognition system in^[6] was designed using data observed by Microsoft Kinect. The system was depended on the evaluation of specific significant tasks of human body. K-means clustering support vector machines and hidden Markov model were integrated for predicting the postures added during the performance of an activity. Then, these activities were categorized for designing each activity as a spatiotemporal progress of recognized postures. The issues related to misclassification were not addressed efficiently.

A Selective Subsequence Time Series (SSTS) clustering in^[7] was described to offer significant outcomes through data encoding to cluster only necessary subsequences. The error rate was not decreased to the desired level by employing SSTS clustering. A smartphone inertial sensors-based approach was implemented in^[8] for human activity detection. The significant features were mined from unprocessed data and these features were processed by using kernel principal component analysis (KPCA) and linear discriminant analysis (LDA) for achieving more efficiency. Then, the features were applied with a Deep Belief Network (DBN) to improve the accuracy of activity detection but failed in considering efficient and complex activity's recognition in real-time environments.

A deep convolutional neural network (convnet) was introduced in^[9] for detecting the human activities in a more efficient manner with the aid of smartphone sensors. The natural features of actions and 1D time-series signals were developed. In addition to, significant features from unrefined data was extracted repeatedly and data-adaptively. Integration of Support Vector Machine (SVM) was not performed to attain higher human activity recognition accuracy.

An integration of WiFi-based and Kinect-based activity recognition system referred as HuAc was developed in^[10] for monitoring human activity in an indoor setting. A subcarrier selection method was implemented with respect to sensitivity of subcarriers to human events. Then, spatial association of adjacent skeleton joints was optimized and respective association between channel state information (CSI) was extracted with activity detection depended on skeleton. The precision of human activity recognition was improved but failed in facilitating the practical application of human activity recognition.

A recognition model of multiple dictionaries of action primitives was developed in^[11] to consider the complementary features of motions. Decoupled fusion of several representations assisted in enhancing classification accuracy. However, Time complexity issues were not resolved.

Suggestion and Future Improvement

The future direction of survival study is to enhance the performance of human activity recognition with high accuracy and lesser time consumption by using advanced machine learning methods.

CONCLUSION

A comparison of different machine learning methods for discovering the significant features during human activity recognition is studied. From the survival study, the detection of discriminative features was not carried out in efficient manner as the computational complexity of human activity recognition was not decreased by means of existing techniques. The error rate occurred during the discovery of human activities was also not reduced to the desired level. A wide range of experiments on existing methods and algorithm evaluates the comparative results of various machine learning techniques with its limitations. Finally, based on the limitations identified from the existing works, the future work which is to be carried out for the improvement of machine learning techniques to enhance the desirable accuracy, high precision with the lesser error rate and speedup the time level in the further research work for attaining an efficient human activity recognition method in extracting discriminative features.

REFERENCES

1. Zhenghua Chen, Qingchang Zhu, Yeng Chai Soh, and Le Zhang, 2017. "Robust Human Activity Recognition Using Smartphone Sensors via CT-PCA and Online SVM", *IEEE Transactions on Industrial Informatics*, **13**(6): 3070 – 3080.
2. Yaqiang Yao, Yan Liu, Zhenyu Liu, Huanhuan Chen, 2018. "Human Activity Recognition with Posture Tendency Descriptors on Action Snippets", *IEEE Transactions on Big Data*, **PP**(99): 1-13.
3. Andrey Ignatov, 2018. "Real-time human activity recognition from accelerometer data using Convolutional Neural Networks", *Applied Soft Computing*, *Elsevier*, **62**: 915-922.
4. Deepika Singh, Erinc Merdivan, Ismini Psychoula, Johannes Kropf, Sten Hanke, Matthieu Geist, and Andreas Holzinger, 2017. "Human Activity Recognition Using Recurrent Neural Networks", *International Federation for Information Processing*, Springer, pp. 267-274.

5. Lina Yao, Quan Z. Sheng, Xue Li, Tao Gu, Mingkui Tan, Xianzhi Wang, Sen Wang, and Wenjie Ruan, 2018. “Compressive Representation for Device-Free Activity Recognition with Passive RFID Signal Strength”, *IEEE Transactions on Mobile Computing*, **17**(2): 293 – 306.
6. Gaglio, S., Re, G. and Morana, M. 2015. “Human activity recognition process using 3-d posture data”, *IEEE Transactions on Human- Machine Systems*, **45**(5): 586–597.
7. Sura Rodpongpun, Vit Niennattrakul, Chotirat Ann Ratanamahatana, 2012. “Selective Subsequence Time Series clustering”, *Knowledge-Based Systems*, Elsevier, **35**: 361-368.
8. Mohammed Mehedi Hassan, Md. Zia Uddin, Amr Mohamed and Ahmad Almogren, 2018. “A robust human activity recognition system using smartphone sensors and deep learning”, *Future Generation Computer Systems*, Elsevier, **81**: 307-313.
9. Charissa Ann Ronao and Sung-Bae Cho, 2016. “Human activity recognition with smart phone sensors using deep learning neural networks”, *Expert Systems with Applications*, Elsevier, **59**: 235-244.
10. Linlin Guo, Lei Wang, Jialin Liu, Wei Zhou, and Bingxian Lu, 2018. “HuAc: Human Activity Recognition Using Crowdsourced WiFi Signals and Skeleton Data”, *Hindawi Wireless Communications and Mobile Computing*, pp. 1-15.
11. Amir, H., Shabani, John, S., Zelek, David A. Clausi, 2013. “Multiple scale-specific representations for improved human action Recognition”, *Pattern Recognition Letters*, Elsevier, **34**(15): 1771-1779.

