International Journal of Advances in Applied Sciences (IJAAS)
Vol. 13, No. 2, June 2024, pp. 429~438
ISSN: 2252-8814, DOI: 10.11591/ijaas.v13.i2.pp429-438

Enhanced human activity recognition through deep multi-layer perceptron on the UCI-HAR dataset

Md. Anwar Hossain[1], Sajeeb Kumar Ray[1], Naima Islam[1], Alamin[1], Mirza AFM Rashidul Hasan[2]
[1]Department of Information and Communication Engineering, Faculty of Engineering and Technology,
Pabna University of Science and Technology, Pabna, Bangladesh
[2]Department of Information and Communication Engineering, Faculty of Engineering, University of Rajshahi, Rajshahi, Bangladesh

Article Info

Article history:

Received Feb 7, 2024
Revised Feb 25, 2024
Accepted Mar 10, 2024

Keywords:

Artificial neural network
Convolutional neural network
Human activity recognition
Multi-layer perceptron
UCI-HAR

ABSTRACT

Using the UCI-HAR dataset, this paper examines human activity recognition (HAR) from the perspectives of data science and artificial intelligence. The primary objective is to present and evaluate the effectiveness of a multi-layer perceptron (MLP) model, concentrating on six different activity categories. We train and assess the MLP model using the UCI-HAR dataset, contrasting its results with those of convolutional neural networks (CNN). The MLP model shows competitive results, attaining an amazing 97% validation and testing accuracy, highlighting its efficiency for smaller datasets. An extensive study is carried out to assess the model's adaptation to a larger Motion Sense dataset using confusion matrices and cross-entropy, the model shows robustness with an accuracy of 89%. The MLP model performs admirably, demonstrating its capacity to pick up complex patterns. Results from comparative analysis with CNN are competitive, especially when dealing with smaller datasets. The suggested MLP model shows up as a practical and efficient way to advance HAR techniques. Its remarkable performance and versatility not only show its usefulness in real-world scenarios but also point to interesting directions for further study in the area of HAR.

Corresponding Author:

Md. Anwar Hossain
Department of Information and Communication Engineering, Faculty of Engineering and Technology
Pabna University of Science and Technology
Dhaka Pabna Hwy, Pabna 6600, Bangladesh
Email: manwar.ice@pust.ac.bd

1. INTRODUCTION

Human activity recognition (HAR) is an area of study and application that is quickly emerging in the domains of artificial intelligence and data science [1]. It involves the automatic identification and classification of human actions based on data provided by sensors [2]. These activities can include a broad range of movements and behaviors, including sitting, running, walking, and even more complex movements like driving or cooking. HAR is required associated with the possibility it has to transform several industries, such as security systems, healthcare, smart homes, fitness tracking, supported living for the elderly, and sporting analysis of performance [3]–[6]. Through precise identification of human actions in either the present or past, HAR provides a significant understanding of human behavior and facilitates tailored interventions, enhanced security measures, and improved user experiences [7]. While the conceptual frameworks of vision-based HAR methods are well-established, certain limitations hinder their practical implementation, such as the influence of ambient lighting, camera placement, potential obstacles, and security concerns [8]. One of the primary approaches in HAR is the utilization of sensor data, often obtained from wearable devices like accelerometers and gyroscopes [9]. These sensors capture intricate patterns of movement and orientation, forming the foundational data that supports HAR models. Due to its ability to provide abundant, continuous, and multidimensional knowledge of human movements, sensor data is the most effective option for HAR as it enables precise and comprehensive activity recognition [10].

The UCI-HAR dataset is widely used for evaluating and training HAR models [11]. This dataset includes accelerometer and gyroscope data from thirty individuals who performed various everyday tasks such as walking, ascending stairs, descending stairs, sitting, standing, and lying down. With its extensive and comprehensive set of labeled data, the UCI-HAR dataset is particularly suitable for research and experimentation purposes, allowing for thorough testing and training of HAR algorithms [12]. The dataset's reliability, inclusiveness, and overall size contribute to its efficiency, ensuring accurate performance evaluation and enabling the development and validation of HAR models with varying levels of complexity. Prior to analysis, preliminary techniques like feature extraction, normalization, and data cleaning are crucial steps in preparing the data [13]. Activities are classified using a range of deep learning and machine learning models, including convolutional neural networks (CNNs) [14], recurrent neural networks (RNNs) [15], and support vector machines (SVMs) [16].

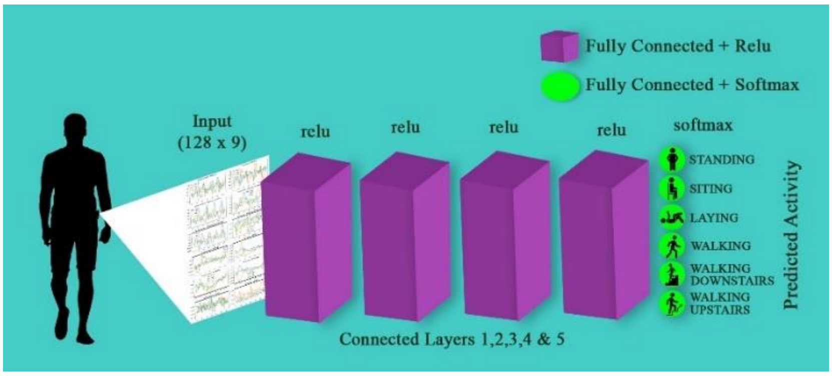
Most previous work on the topic of human activity classification has concentrated on identifying human activities using a range of sensor data. Models based on CNN suggested tremendous potential, superior performance other state-of-the-art techniques in terms of recognition accuracy. Nair *et al*. [17] employed the Instantaneous CNN architecture, a class of contemporary models utilizing a hierarchical framework of temporal convolutions. According to Dua *et al*. [18] employ CNNs, an end-to-end model known as the Gated Recurrent Unit is proposed that can automatically extract features and categorize the actions. A CNN-based approach to sensor fusion was proposed by Münzner *et al*. [19] to overcome difficulties in normalizing and combining data from multimodal sensors. Nevertheless, neither approach could discern between comparable actions such as standing and sitting. Researchers in the field of HAR have also applied RNN, a deep learning technique, extensively. The ability of RNNs to make sense of spatial data sequences is one of its specific characteristics. For example, wearable or inertial sensor-based HAR greatly benefits from long short-term memory (LSTM)-based networks, which can capture long-term relationships within data sequences. Agarwal and Alam [20] developed a lightweight model for activity recognition using LSTM and shallow RNN. However, the accuracy of their model did not meet the required standards. Research by Zebin *et al*. [21] was able to classify human activities by capturing spatiotemporal characteristics through the use of LSTM-based architectures, where they faced difficulties in accurately distinguishing between walking downstairs and walking, as well as sitting and standing. Hybrid models, such as CNN-RNN [22], CNN-LSTM [23], and CNN-gated recurrent unit (GRU) [24], were also presented in several studies, and they significantly increased the accuracy of identification. However, when convolutional layers and other methods are integrated, the computational complexity increases and results in a larger number of hyperparameters.

Our research endeavors are concentrated on addressing the gaps in knowledge present within the UCI-HAR dataset, with the ultimate goal of creating highly efficient and instantly applicable models for HAR. Our compact deep-learning model achieves 96.81% accuracy in under three minutes of training. Ideal for resource-constrained devices, like microcontrollers, smartphones, and smartwatches. Combining traditional ML with deep learning techniques enhances identification precision and extracts more data points. Our suggested multi-layer perceptron (MLP) model, prioritizes simplicity for faster execution and outperforms existing deep cognition models. Anticipating further improvements, our model represents a significant advancement in HAR research.

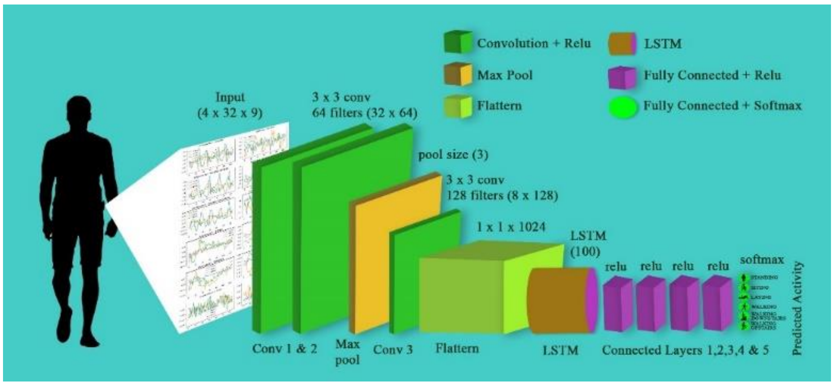
2. THE PROPOSED METHOD

2.1. Architectural design

The system architectures are shown in Figure 1, where Figure 1(a) presents our proposed model architecture, replacing the intricate CNN architecture showcased in Figure 1(b). Our proposed model's neural network is built using the Sequential API, which suggests a progressive layer-by-layer framework. At the forefront lies a dense layer boasting 64 units, employing the activation function known as the rectified linear unit (ReLU), accompanied by an ordinary weight initialization method. The input dimension of this layer is carefully chosen, taking into account the nature of the training data. Subsequently, a dropout layer enters the scene, randomly discarding 20% of the units in each training batch, serving as a safeguard against the perils of overfitting. The subsequent layers consist of a dense layer with 36 units, another with 12 units, yet another with 36 units, and finally, a concluding layer housing 6 units. Except for the final layer, which embraces the softmax activation function, all these layers employ the ReLU activation function and normal weight initialization.



(a)



(b)

Figure 1. System architecture (a) proposed MLP architecture for the identification of human activity and (b) CNN architecture for identifying human activity (conv refers to a convolutional layer)

2.2. Algorithm of the proposed model

The CNN algorithm for HAR is explained in Algorithm 1. It describes the steps used to create the CNN model, which includes the convolutional layers, LSTM layer, max pooling, batch normalization, dropout layers, and dense layers with ReLU activation functions. The suggested method for HAR is shown in Algorithm 2. It explains a sequential model for classification that has a dropout layer, dense layers, and a final dense layer with a softmax activation function.

Algorithm 1. CNN algorithm for HAR

Initialize the model	Model=Sequential ()
Convolution1D →	Conv1D (filter, kernel dimensions, padding, input shape, activation)
	Batch Normalization ()
Convolution1D →	Conv1D (filter, kernel dimensions, activation, padding)
Max pooling →	MaxPooling1D (pool dimensions)
	Dropout (rate)
Convolution1D →	Conv2D (filter, kernel dimensions, activation, padding)
	Batch Normalization ()
	Flatten ()
LSTM →	LSTM (units)
	Dropout (rate)
Dense →	Dense (Units, Karnel initializer, activation=”Relu”)
	Dropout (rate)
Dense →	Dense (Units, Karnel initializer, activation=”Relu”)
Dense →	Dense (Units, Karnel initializer, activation=“Relu”)
Dense →	Dense (Units, Karnel initializer, activation=“Relu”)
Dense →	Dense (Total number of categorization classes, activation="softmax")
End	

Algorithm 2. Proposed algorithm for HAR

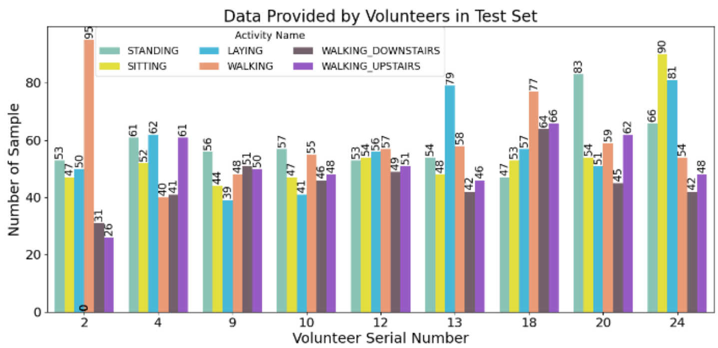
Initialize the model		Model=Sequential ()
1	Dense →	Dense (Units, Karnel initializer, activation=”Relu”, input shape)
2		Dropout (rate)
3	Dense →	Dense (Units, Karnel initializer, activation=”Relu”)
4	Dense →	Dense (Units, Karnel initializer, activation=“Relu”)
5	Dense →	Dense (Units, Karnel initializer, activation=“Relu”)
6	Dense →	Dense (Total number of categorization classes, activation="softmax")
End		

3. RESEARCH METHOD

3.1. Dataset ingestion and initial analysis

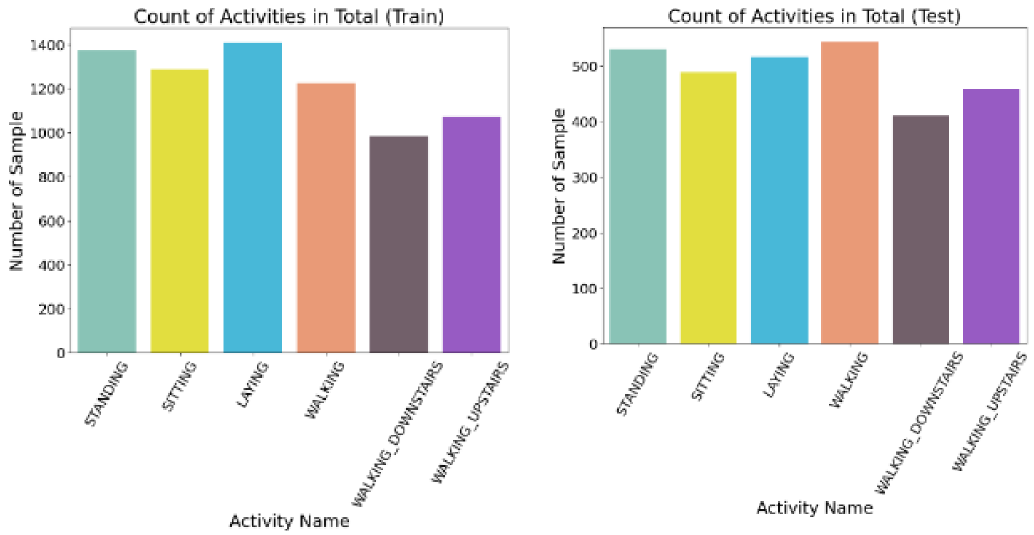
For further study, the UCI-HAR dataset was acquired and imported into a Pandas DataFrame. The dataset contains six measurements from sensors namely gyroscope and accelerometer. A Label Encoder was utilized to convert the activity labels into numerical values. To standardize the sensor data and mitigate potential biases arising from variations in feature sizes, feature scaling was performed. The study included a total of thirty participants. The dataset was divided into a training dataset and a test dataset to train and evaluate the model. Specifically, the model was trained using data from 21 volunteers, and its performance was evaluated using the test dataset, which contained data from the remaining 9 individuals. Graphical representations were created to visually depict the distribution of activities within these databases. Figure 2 presents a visual depiction of the characteristics of the entire dataset, showcasing the patterns and variations evident in the collected data. Where Figure 2(a) represents the training dataset and Figure 2(b) represents the test dataset. Additionally, a comprehensive analysis is displayed in Figure 3 to comprehend the structure of the tasks performed in the training and test datasets. Where Figure 3(a) represents the training dataset and Figure 3(b) represents the test dataset. These visualizations provide insights into the distribution of activities and their relative frequencies in the training and test datasets by showcasing the count of instances for each unique activity.

(a)



(b)

Figure 2. Overview of volunteer demographics in (a) train dataset and (b) test dataset



(a) (b)

Figure 3. Distribution of activities in (a) train dataset and (b) test dataset

3.2. Sensor signals visualization

The signals were sampled by employing 128 readings for each window of sliding windows with a fixed width, incorporating a 50% overlap and lasting 2.56 seconds. The sampled signal visualization is depicted in Figure 4. This temporal segmentation greatly aided in the identification of significant patterns and characteristics within the time-series data shown in Figure 4(a). One of the most important steps in HAR is the conversion of time-domain signals to frequency domains, enabling a more thorough analysis. This conversion is achieved through the utilization of a signal manipulation technique known as fast Fourier transform (FFT). Notably, the frequency domain representations, as demonstrated in Figure 4(b), offer valuable insights into the underlying patterns associated with diverse activities.

3.3. Compilation of the model

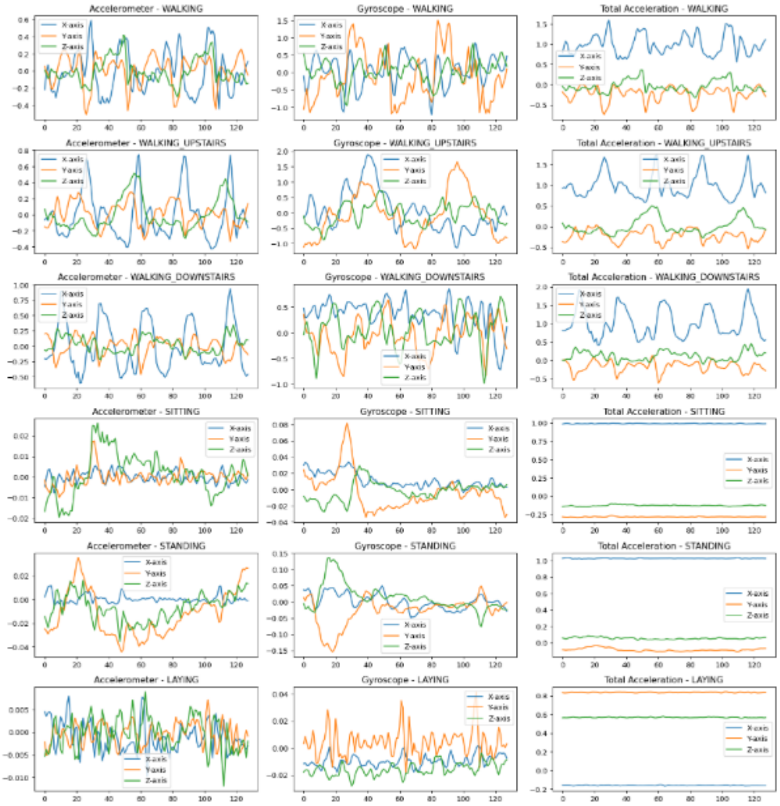
The model is made ready for training during the process of compilation by establishing three crucial elements: loss function (SparseCategoricalCrossentropy): this loss function is used to quantify, during training, the discrepancy between the true labels and the model's predictions. SparseCategorical-Crossentropy is used for multi-class classification tasks, where each input corresponds to a single class. Optimizer (Adam): the optimizer determines how to update the weights of the model based on the determined slopes of the loss function. Adam is a popular optimizer that adjusts the learning rate dynamically to increase the efficiency of weight updates during training. Learning rate (0.001): this hyperparameter controls how big of a step the optimizer takes when changing the model's weights. In this model, it remains fixed at 0.001.

3.4. Model training procedure

Throughout the training phase, the trained model receives input from training data (x_train and y_train), and iterative updates to the model's parameters (weights and biases) are made to minimize the given loss function. Key training variables include Batch size (64): the batch size is the number of samples processed during a single training phase. The parameters of the model are changed once for every batch. Epochs (250): the total number of times the neural network runs the training dataset both forward and backward. There are multiple data batches in each epoch. Validation data (x_test, y_test): an impartial subset of the dataset that's meant to detect overfitting and assess how well the model performs on missing data.

3.5. Model prediction procedure

After training, the model is applied to the test data (x_test) for prediction. The actions to take are as follows: Logits calculation (logits=model (x_test)): the logits, or unnormalized predictions produced by the model, are computed by applying the model to the test data. Softmax activation (prediction=tf.nn.softmax (logits)): the predicted probabilities for every class are obtained by applying Softmax activation, which guarantees that they add up to 1 and reflect the confidence of the model for every class. Predicted classes (pred_class=np.argmax (prediction, axis=1)): The category with the highest probability is selected to determine the predicted class for each input sample using argmax along the appropriate axis.



(a)

(b)

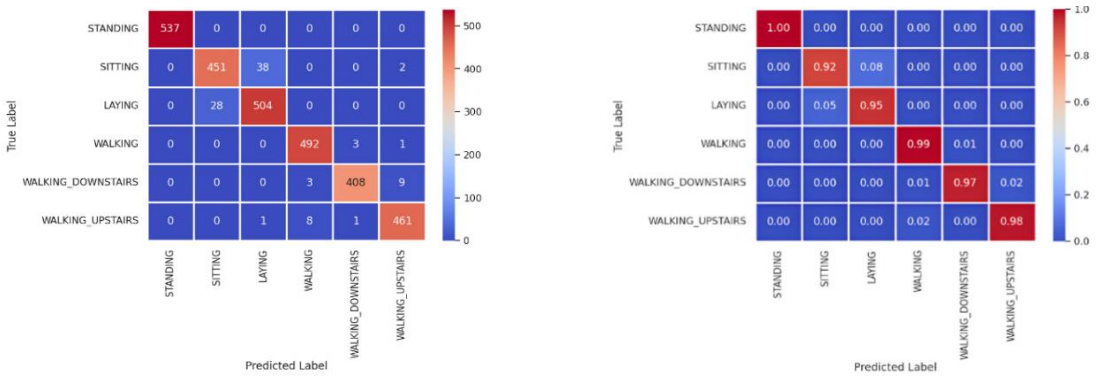
Figure 4. Sampled signal visualization of different activity in (a) time domain representation and (b) frequency domain representation

4. RESULTS AND DISCUSSION

In our research, our objective was to address significant gaps in the field of HAR. To achieve this, we introduced a specialized MLP model for the UCI-HAR dataset. Previous research lacked efficient models that were suitable for real-time applications. Our objective was to fill this gap and provide a lightweight yet precise solution. The proposed model achieved an accuracy of 97% in testing and validation. The training process utilized the Adam optimizer, a dropout rate of 20%, and a learning rate of 0.001. Throughout 250 epochs, we meticulously monitored the accuracy and loss trends for both the training and validation sets. We examine the effectiveness of our proposed model further by using a confusion matrix and a normalized confusion matrix. The confusion matrix compared projected outputs to actual outputs. The normalized confusion matrix scaled values between 0 and 1, with 1 representing 100% accuracy.

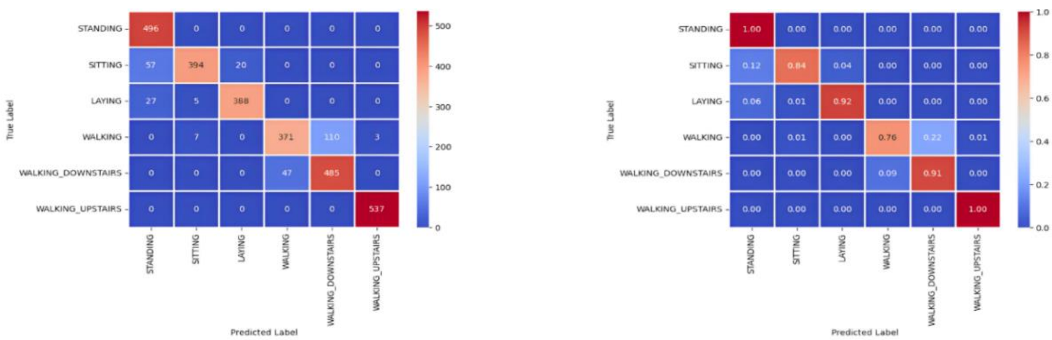
Figure 5 demonstrates the assessment of the effectiveness of our proposed model. The confusion matrix is showcased in Figure 5(a), while Figure 5(b) presents the corresponding normalized confusion matrix. However, when we applied our CNN model to the UCI-HAR dataset, its performance experienced a significant decline. Despite having a training accuracy of 97%, the CNN's validation and testing accuracies were only 91%. To monitor accuracy and loss trends over 200 epochs, we employed the Adam optimizer with a 20% dropout rate and a learning rate of 0.00001. Figure 6 displays the evaluation of the CNN model. Where, Figure 6(a) depicts the confusion matrix, and Figure 6(b) exhibits the normalized confusion matrix. Based on this comparison, we conclude that the proposed MLP model is better suited for the UCI-HAR dataset than the CNN architecture.

Furthermore, our modified MLP model exhibited impressive performance on the MotionSense dataset, achieving evaluation and verification accuracies of 89%. This improvement was achieved by increasing the number of neurons in the current linked layers, using an Adam optimizer, a dropout rate of 20%, and a learning rate of 0.001. The performance evaluation of this is shown in Figure 7. Figure 7(a) presents the confusion matrix, while Figure 7(b) presents the normalized confusion matrix. These visuals highlight the model's commendable performance even when dealing with a larger dataset.



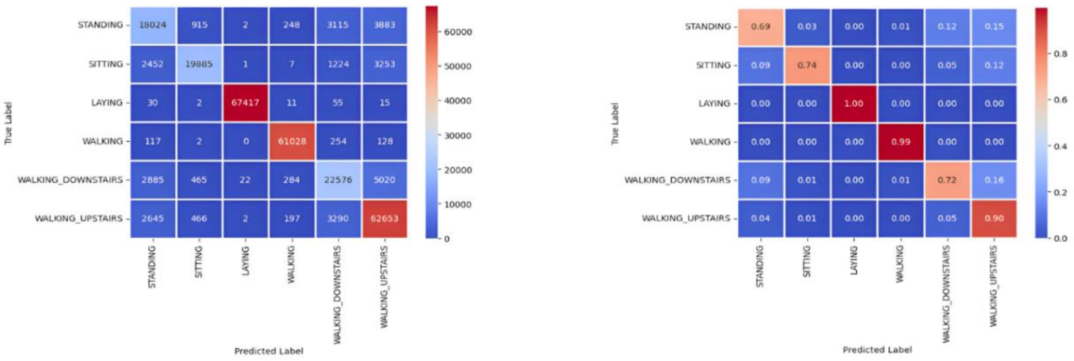
(a) (b)

Figure 5. UCI-HAR dataset confusion matrix and corresponding normalized confusion matrix using MLP (a) confusion matrix and (b) normalized confusion matrix



(a) (b)

Figure 6. UCI-HAR dataset confusion matrix and corresponding normalized confusion matrix using CNN (a) confusion matrix and (b) normalized confusion matrix



(a) (b)

Figure 7. MosionSense dataset confusion matrix and corresponding normalized confusion matrix using MLP (a) confusion matrix and (b) normalized confusion matrix

Table 1 illustrates how our proposed MLP model fared compared to other methods on the UCI-HAR dataset. Specifically, our model performs better than the CNN-LSTM model that was published in [23], which only managed to get an accuracy rate of 92.13%. With an accuracy of 96.71%, CNN-GRU [24] proved effective; nevertheless, because of its computational complexity, which arises from the combination of recurrent units and convolutional layers, a greater number of hyperparameters are needed. The accuracy rates of other techniques, such as the CNN-LSTM [25] and CNN-SVM [26] architectures, were 91.89% and 95% respectively. Additionally, a CNN method developed by [27] achieved a 95.25% accuracy rate.

Table 1. Comparison of the proposed model with the CNN approach

Method	Model	Dataset	Accuracy (%)
[23]	CNN-LSTM	UCI-HAR	92.13
[24]	CNN-GRU	UCI-HAR	96.71
[25]	CNN-LSTM	UCI-HAR	91.89
[26]	CNN-SVM	UCI-HAR	95
[27]	CNN	UCI-HAR	95.25
Our	CNN	UCI-HAR	90.63
Proposed	MLP	*MotionSense*	89.03
method	**MLP**	**UCI-HAR**	**96.81**

Although our work sheds light on HAR, there are several limitations. Our study was mainly conducted on the UCI-HAR and MotionSense datasets, for example, which would have limited the applicability of our results to other datasets or real-world situations. To further investigate the scalability and applicability of our suggested MLP model to bigger and more varied datasets, more research may be required. Our results provide new avenues for future HAR studies. In particular, to enhance the suggested MLP model's performance across a range of datasets and in practical applications, subsequent research might focus on fine-tuning and optimizing it. Further research on workable approaches to HAR system implementation in many scenarios may yield important new understandings of the applicability and usefulness of these systems. The findings of this research suggest that MLP offers a faster and more accurate alternative for small datasets like UCI-HAR compared to CNNs. Additionally, the MLP model that we propose exhibits extraordinary accuracy when tested and validated using the MotionSense dataset. This achievement is particularly noteworthy due to the MotionSense dataset's larger size as compared to UCI-HAR, thereby signifying the robustness and efficiency of the MLP model.

5. CONCLUSION

We propose a strong MLP model for HAR using the UCI-HAR dataset. Although the proposed model is simple, it competes well with larger models like CNN. In fact, our model outperforms CNN in terms of accuracy and provides a faster solution for smaller datasets. We emphasize the model's ability to effectively learn complex patterns and its practicality through the use of confusion matrices. Additionally, the model's exceptional results on the MotionSense dataset suggest that it may be scalable to larger datasets. We firmly believe that our approach has the potential to significantly advance the field of human activity detection and analysis on smaller devices. Moving forward, our goal is to continuously enhance our algorithm's capability to recognize a wide range of human actions.

REFERENCES

[1] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6M: large scale datasets and predictive methods for 3D human sensing in natural environments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 7, pp. 1325–1339, Jul. 2014, doi: 10.1109/TPAMI.2013.248.

[2] A. Jain and V. Kanhangad, "Human activity classification in smartphones using accelerometer and gyroscope sensors," *IEEE Sensors Journal*, vol. 18, no. 3, pp. 1169–1177, Feb. 2018, doi: 10.1109/JSEN.2017.2782492.

[3] A. Zahin, L. T. Tan, and R. Q. Hu, "Sensor-based human activity recognition for smart healthcare: a semi-supervised machine learning," in *Artificial Intelligence for Communications and Networks. AICON 2019. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Cham: Springer, 2019, pp. 450–472. doi: 10.1007/978-3-030-22971-9_39.

[4] J. Manjarres, P. Narvaez, K. Gasser, W. Percybrooks, and M. Pardo, "Physical workload tracking using human activity recognition with wearable devices," *Sensors*, vol. 20, no. 1, p. 39, Dec. 2019, doi: 10.3390/s20010039.

[5] Du, Lim, and Tan, "A novel human activity recognition and prediction in smart home based on interaction," *Sensors*, vol. 19, no. 20, p. 4474, Oct. 2019, doi: 10.3390/s19204474.

[6] M. M. Hassan, M. Z. Uddin, A. Mohamed, and A. Almogren, "A robust human activity recognition system using smartphone sensors and deep learning," *Future Generation Computer Systems*, vol. 81, pp. 307–313, Apr. 2018, doi: 10.1016/j.future.2017.11.029.

[7] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013, doi: 10.1109/SURV.2012.110112.00192.

[8] A. Sarkar, S. K. S. Hossain, and R. Sarkar, "Human activity recognition from sensor data using spatial attention-aided CNN with genetic algorithm," *Neural Computing and Applications*, vol. 35, no. 7, pp. 5165–5191, Mar. 2023, doi: 10.1007/s00521-022-07911-0.

[9] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, Mar. 2011, doi: 10.1145/1964897.1964918.

[10] G. Diraco, G. Rescio, P. Siciliano, and A. Leone, "Review on human action recognition in smart living: sensing technology, multimodality, real-time processing, interoperability, and resource-constrained processing," *Sensors*, vol. 23, no. 11, p. 5281, Jun. 2023, doi: 10.3390/s23115281.

[11] J. Reyes-Ortiz, D. Anguita, A. Ghio, L. Oneto, and X. Parra, "Human activity recognition using smartphones." 2012. doi: 10.24432/C54S4K.

[12] S. Khare, S. Sarkar, and M. Totaro, "Comparison of sensor-based datasets for human activity recognition in wearable IoT," in *2020 IEEE 6th World Forum on Internet of Things (WF-IoT)*, IEEE, Jun. 2020, pp. 1–6. doi: 10.1109/WF-IoT48130.2020.9221408.

[13] D. Figo, P. C. Diniz, D. R. Ferreira, and J. M. P. Cardoso, "Preprocessing techniques for context recognition from accelerometer data," *Personal and Ubiquitous Computing*, vol. 14, no. 7, pp. 645–662, Oct. 2010, doi: 10.1007/s00779-010-0293-9.

[14] M. Zeng *et al.*, "Convolutional neural networks for human activity recognition using mobile sensors," in *Proceedings of the 6th International Conference on Mobile Computing, Applications and Services*, ICST, 2014. doi: 10.4108/icst.mobicase.2014.257786.

[15] S. W. Pienaar and R. Malekian, "Human activity recognition using LSTM-RNN deep neural network architecture," May 2019, [Online]. Available: http://arxiv.org/abs/1905.00599

[16] D. Thakur and S. Biswas, "Smartphone based human activity monitoring and recognition using ML and DL: a comprehensive survey," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 11, pp. 5433–5444, Nov. 2020, doi: 10.1007/s12652-020-01899-y.

[17] N. Nair, C. Thomas, and D. B. Jayagopi, "Human activity recognition using temporal convolutional network," in *Proceedings of the 5th International Workshop on Sensor-based Activity Recognition and Interaction*, New York, NY, USA: ACM, Sep. 2018, pp. 1–8. doi: 10.1145/3266157.3266221.

[18] N. Dua, S. N. Singh, and V. B. Semwal, "Multi-input CNN-GRU based human activity recognition using wearable sensors," *Computing*, vol. 103, no. 7, pp. 1461–1478, Jul. 2021, doi: 10.1007/s00607-021-00928-8.

[19] S. Münzner, P. Schmidt, A. Reiss, M. Hanselmann, R. Stiefelhagen, and R. Dürichen, "CNN-based sensor fusion techniques for multimodal human activity recognition," in *Proceedings of the 2017 ACM International Symposium on Wearable Computers*, New York, NY, USA: ACM, Sep. 2017, pp. 158–165. doi: 10.1145/3123021.3123046.

[20] P. Agarwal and M. Alam, "A lightweight deep learning model for human activity recognition on edge devices," *Procedia Computer Science*, vol. 167, pp. 2364–2373, 2020, doi: 10.1016/j.procs.2020.03.289.

[21] T. Zebin, M. Sperrin, N. Peek, and A. J. Casson, "Human activity recognition from inertial sensor time-series using batch normalized deep LSTM recurrent networks," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, Jul. 2018, pp. 1–4. doi: 10.1109/EMBC.2018.8513115.

[22] M. Lv, W. Xu, and T. Chen, "A hybrid deep convolutional and recurrent neural network for complex activity recognition using multimodal sensors," *Neurocomputing*, vol. 362, pp. 33–40, Oct. 2019, doi: 10.1016/j.neucom.2019.06.051.

[23] R. Mutegeki and D. S. Han, "A CNN-LSTM approach to human activity recognition," in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, IEEE, Feb. 2020, pp. 362–366. doi: 10.1109/ICAIIC48513.2020.9065078.

[24] C. Zhang, K. Cao, L. Lu, and T. Deng, "A multi-scale feature extraction fusion model for human activity recognition," *Scientific Reports*, vol. 12, no. 1, p. 20620, Nov. 2022, doi: 10.1038/s41598-022-24887-y.

[25] M. F. Aktter, M. A. Hossain, S. Sarker, A. Z. Abadin, and M. A. R. Hasan, "Performance Analysis of deep learning based human activity recognition methods," *Journal of Applied Science & Process Engineering*, vol. 9, no. 2, pp. 1197–1208, Oct. 2022, doi: 10.33736/jaspe.4639.2022.

[26] A. A. Nithya, K. Ishwarya, G. Mummaneni, and V. Verma, "CNN based identifying human activity using smartphone sensors," in *2022 International Conference on Edge Computing and Applications (ICECAA)*, IEEE, Oct. 2022, pp. 1115–1120. doi: 10.1109/ICECAA55415.2022.9936202.

[27] N. Sikder, M. S. Chowdhury, A. S. M. Arif, and A.-A. Nahid, "Human activity recognition using multichannel convolutional neural network," in *2019 5th International Conference on Advances in Electrical Engineering (ICAEE)*, IEEE, Sep. 2019, pp. 560–565. doi: 10.1109/ICAEE48663.2019.8975649.

BIOGRAPHIES OF AUTHORS

Md. Anwar Hossain received B.Sc. (Honours) and M.Sc. degrees in Information and Communication Engineering from the University of Rajshahi, Bangladesh in 2005 and 2006 respectively. He received his M.Phil. degree from the Pabna University of Science and Technology, Bangladesh in 2020. In 2010 he served as a Lecturer in the Department of Information and Communication Technology of Comilla University, Bangladesh. In 2012, he joined Pabna University of Science and Technology, Bangladesh as a faculty member, where he is currently serving as an Associate Professor in the Department of Information and Communication Engineering. Now, he is a Ph.D. student in the Department of Information and Communication Engineering (ICE), Pabna University of Science and Technology (PUST), Bangladesh. His research interests include deep learning, machine learning, image classification, and natural language processing. He can be contacted at email: manwar.ice@gmail.com.

Sajeeb Kumar Ray is currently a bachelor's student in the Department of Information and Communication Engineering (ICE) at Pabna University of Science and Technology (PUST), Bangladesh. His research interests include artificial intelligence, machine learning, deep learning, and image processing. He can be contacted at email: sajeebbro2002@gmail.com.

Naima Islam is currently a bachelor's student in the Department of Information and Communication Engineering (ICE) at Pabna University of Science and Technology (PUST), Bangladesh. Her research interests include artificial intelligence, machine learning, deep learning, and image processing. She can be contacted at email: naimaislam204100@gmail.com.

Alamin is pursuing a bachelor's degree in Information and Communication Engineering at Pabna University of Science and Technology in Bangladesh. His academic pursuits revolve around the Internet of Things (IoT), digital signal processing, deep learning, and image processing. He can be contacted at email: alaminabl20110@gmail.com.

Mirza AFM Rashidul Hasan graduated from the University of Rajshahi in Bangladesh with a B.Sc. (Hons), an M.Sc., and an M.Phil. in applied physics and electronic engineering in 1992, 1993, and 2001, respectively. He began teaching at the University of Rajshahi in Bangladesh in 2006, and he presently holds the position of Professor in the Department of Information and Communication Engineering. From 2003 to 2004, he served as a guest researcher at Waseda University in Japan. From 2006 to 2007, he served as a junior fellow at IWMI. He earned his Ph.D. from the Islamic University's Faculty of Applied Science and Technology in Kushtia, Bangladesh, in 2009, and his D.Engg. from the Graduate School of Science and Engineering at Saitama University in Saitama, Japan, in 2012. His current areas of interest are voice, picture, and communication systems applications of digital signal processing. He can be contacted at email: mirzahasanice@gmail.com.