Enhanced Human Activity Recognition using Inertial Sensor Data from Smart Wearables: A Neural Network Approach with Residual Connections

Shaida Muhammad
Department of Computing
National University of Sciences & Technology (NUST)
Islamabad, Pakistan
smuhammad.mscs19seecs@seecs.edu.pk

Kiran Hamza
Department of Computing
National University of Sciences & Technology (NUST)
Islamabad, Pakistan
kmehmood.mscs19seecs@seecs.edu.pk

Hamza Ali Imran
Generative AI Group
Emumba Private Limited
Islamabad, Pakistan
himran.mscs18seecs@seecs.edu.pk

Ataul Aziz Ikram
Department of Electrical Engineering
National University of Computer & Emerging Sciences
Islamabad, Pakistan
ata.ikram@nu.edu.pk

Saad Wazir
School of Computing
Korea Advanced Institute of Science & Technology (KAIST)
Daejeon, South Korea
saad.wazir@kaist.ac.kr

***Abstract*—Human Activity Recognition (HAR) has gained significant interest in various research circles due to its wide-ranging applications, including patient monitoring, gaming, and education. While computer vision has traditionally played a key role in HAR, it encounters challenges like privacy concerns and environmental factors such as occlusion. To address these issues, inertial sensors like accelerometers and gyroscopic sensors have gained popularity, offering advantages such as cost-effectiveness and increased mobility. In this study, we propose a Convolutional Neural Network (CNN) architecture for HAR that incorporates residual connections. We rigorously evaluate our model using a freely available dataset WISDM (2011) from the wireless sensor data mining lab and compare its performance with state-of-the-art techniques. Our results indicate that our proposed model not only outperforms existing methods in terms of accuracy but also exhibits reduced complexity, requiring only 38,342 trainable parameters. Our HAR model achieves an impressive average accuracy of 98.32%, along with notable F1-score, Recall, and Precision values of 97.50%. These findings underscore the efficacy of our approach in advancing HAR technology while maintaining efficiency.**

I. Introduction

The reliable recognition of human activities via smart devices equipped with wearable sensors is critical across a wide range of applications, including health surveillance, sports, and entertainment, as well as machine-computer interfaces, childcare, online training, and even the abolition of harmful habits [1], [2]. As a result, HAR has emerged as a pivotal research domain. In the literature, three approaches to HAR are commonly used: computer vision, environment sensors, and on-body sensors, the latter of which include accelerometers, gyroscopes, and heart rate sensors.

However, computer vision-based HAR solutions have several limitations, including privacy concerns, higher operational costs, and susceptibility to environmental variables such as occlusion. Environment-sensor-based HAR solutions constrain the human's location, making them unreliable for long-term monitoring throughout the day in a variety of settings. On-body sensor-based HAR solutions, on the other hand, emerge as the most portable option. The use of inertial sensors, which include accelerometers and gyroscopes, is gaining traction in Human Activity Recognition.

The remarkable growth of micro-electromechanical systems (MEMSs) has given rise to significant advances across various research domains. In this environment, inertial sensors, which form inertia measuring units (IMUs), stand out as critical MEMS sensors. These sensors are appealing because of their advantageous characteristics, which include low cost, low power consumption, lightweight design, and inherent portability. As a result, the use of inertial sensors has increased, either directly or indirectly, across a wide range of applications. The

979-8-3503-4971-9/24/$31.00 © 2024 IEEE

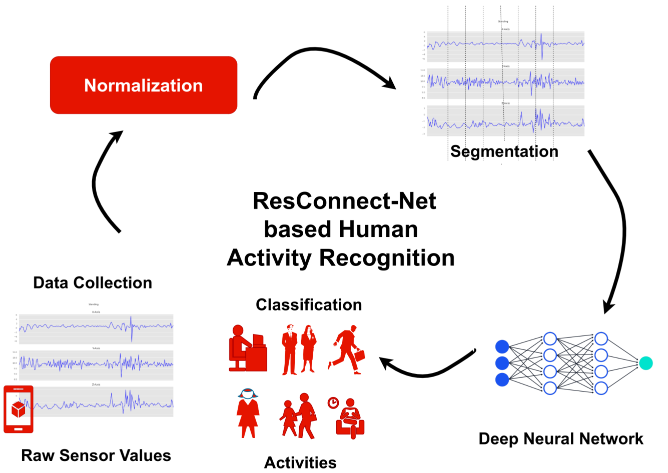

Fig. 1. The entire pipeline of the project: Raw signals are normalized, then segmented before being used for training and inference by the HARResConnect-Net model.

data obtained directly from these sensors are readily available for processing and use in complex motion analysis, made possible by the use of deep learning models [3].

In this article, we introduce a novel convolutional neural network (CNN) which we named as HARResConnect-Net for classifying human activities based on inertial sensor data of smartphones. The presented model contains residual connections. The evaluation of this proposed model is conducted on the WISDM (2011) dataset. The WISDM is a dataset that has been employed in numerous studies to assess Human Activity Recognition (HAR) models. In comparison to prior research, this HARResConnect-Net stands out by surpassing the state-of-the-art. To conduct our assessments, we thoughtfully partitioned the dataset into a ratio of 70:15:15 respectively into training, validation, and test sets. The contributions of our work can be condensed into the following:

- We have introduced an innovative Convolutional Neural Network (CNN) model, distinguished by the inclusion of residual connections, specifically designed to operate directly on inertial sensor data for HAR in raw form.
- We evaluated our model using the WISDM 2011 dataset and benchmarked its performance against prior works.
- Remarkably, our presented model boasts a minimal 38,342 trainable parameters, exemplifying its inherent simplicity and efficiency.

The remainder of the article has the following sections: Section III discusses the proposed methodology, which includes details of the proposed architecture, its motivation and complexity, and the dataset used for the evaluation of the model. Section IV covers the results. Section V provides a conclusion of the work by comparison of the presented model with previous studies.

II. Related Work

Inertial sensors have grown into an important enabling technology in a variety of applications. In the past, they were often employed in navigation, localization, and mapping [4], [5]. With the continuous progress in technology, inertial sensors have become readily accessible through multiple wearable digital gadgets that include smart fitness bands and smartwatches. The utilization of inertial sensors for applications related to human motion is increasingly prevalent. An array of datasets designed for Human Activity Recognition (HAR) leverages these inertial sensors found in wearable devices. Some notable examples encompass datasets such as those by [6]–[10]. These resources contribute significantly to the research and application of inertial sensor data in HAR.

The research presented by [11] delves into the domain of HAR by utilizing data from smartphone inertial sensors. This study involves extensive feature engineering, where various statistical metrics, for example, mean and median are calculated from the raw sensor data. To enhance the robustness of the features, the authors employed techniques such as Kernel Principal Component Analysis. The authors also employed Linear Discriminant Analysis (LDA) along with the KPCA. The derived robust features are subsequently employed to train a Deep Belief Network (DBN). In their study, the proposed approach is rigorously compared with conventional methods, including the multi-class SVM and Artificial Neural Network. The results are noteworthy as the DBN model outperforms these traditional methods, achieving an accuracy rate of 89.61% in classifying 11 distinct activity classes. This achievement highlights the effectiveness of their approach in the realm of HAR and its potential for real-world applications.

Apart from determining activities, IMUs of smart devices are being used for other similar applications as well. The age of a person is predicted in [12] from the walking style called gait using machine and deep learning on IMU of a smartphone phone. Similarly, authors in [13] estimate not just age but also the gender, and height of the person by making use of IMU data of human gait. In [14], the person is re-identified from inertial sensors data of human gait. For this purpose, a deep learning model having bidirectional gated recurrent units is proposed. A study in [15] classified 6 different human emotions using inertial sensor data from a smartphone. They collected data from around 40 volunteers. Volunteers were asked to recall events from the past to enter a particular emotional state and then they were asked to walk. Feature engineering was performed, and the team was capable of classifying 6 emotions with an accuracy of 86%. A similar work is carried out in [16] with 3 different emotions. Hashmi et. al. [17] tries to identify. In [18], a HAR system utilizing a wearable sensor and a deep learning-based neural network was presented. Initially, LSTM and CNN were considered, with CNN ultimately selected for its 2% higher accuracy over LSTM. Optimization of the CNN model led to the selection of the network structure with the highest accuracy. Further testing compared the model's performance using datasets from both the UCI machine learning repository and the proposed sensor, demonstrating exceptional accuracy of 97% across nine distinct activities. These results underscore the model's robustness and potential for real-world applications. In [19], a new

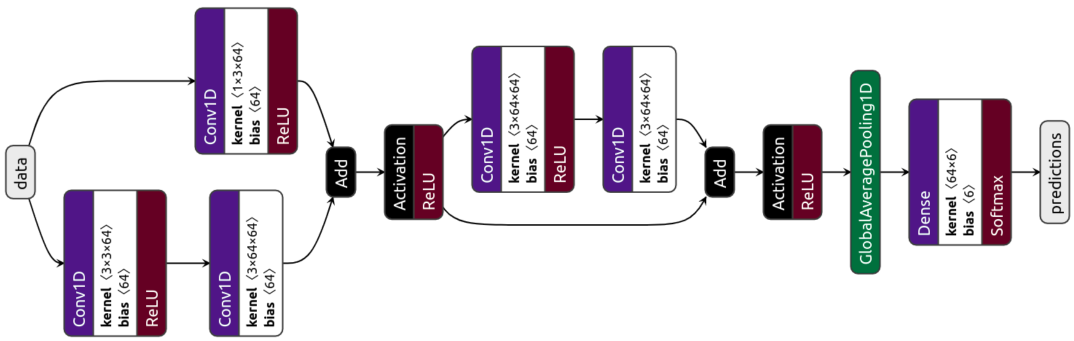

Fig. 2. HARResConnect-Net: The Proposed Model

feature for bidirectional long-short-term memory (BLSTM) in mobile device-based human activity recognition systems (MARS) is proposed, alongside an ensemble classifier for improved performance. Acceleration is divided into horizontal and vertical components to create a novel two-directional feature. The study utilizes the NMHA dataset, collected from 100 subjects via smartphone accelerometers. Experimental results compare the error rates of five algorithms, including SVM, EMR, KNN, BLSTM, and MBLSTM, revealing MBLSTM's superiority with the lowest error rate across all 8 classes. In [20], a multi-task model was proposed for activity classification and intensity estimation using wearable sensor data. Long short-term memory (LSTM) was chosen for both tasks, optimized with Adam Optimizer. For activity classification, a single-layer LSTM with a size of 64 achieved the best accuracy (0.9776) on the test set, with a test subject F1 score of 0.8343. For intensity classification, a two-layer LSTM model with a size of 64 yielded the best results. The integrated Multi-task model achieved an average F1 score of 0.8132.

In [21] design of a human activity classification system using frequency-modulated continuous pulse(FMCW) radar has been proposed. The data set for the experiment was collected using 15 participants and the data set comprises of 6 different human activities. Firstly Fourier transform was applied to raw data from the radar. Static clutters were removed using a notch filter. Doppler time pattern was obtained using a short-time Fourier transform. Three different classifiers were chosen and experiments were conducted. Experiment results showed that Bi-LSTM obtained the best accuracy of 98% so Bi-LSTM was chosen as the deep learning network. Experimental results showed that Doppler domain data using the Bi-LSTM network achieved 90% accuracy whereas by using range domain data only an accuracy of 76% was achieved.

HARResConnect-Net is superior to them previously presented models in terms of accuracy and performance we have achieved 98.32% accuracy and 97.5% F1-Score, Recall and Precision a detailed comparison is available in conclusion section V.

III. Methodology

The overall pipeline of our proposed approach is presented in Figure 1. Our method takes raw sensor values as inputs which are normalized before segmentation. No explicit feature engineering was employed before segmentation making the proposed approach entirely a deep learning one. A novel one-dimensional (1D) convolutional neural network with residual connections is proposed in this work, termed HARResConnect-Net. The data normalization and segmentation steps are common to both the training and test phases.

Normalized sensor data for a sensor axis is obtained by dividing its raw sensor values by its maximum value. To feed the normalized sensor data to the proposed network, it must be segmented. The size of the segmentation window and overlapping between two samples are critical parameters to discover important temporal patterns as highlighted in [1], [22]. In our implementation, the window size is empirically selected to be 128 with 50% overlapping between the two samples for best performance. Consequently, the next sample is taken after shifting the segmentation window 64 entries ahead.

A. Proposed Architecture

The key features in the design of the proposed architecture (as depicted in Figure 2 are multiple kernel sizes that capture different temporal resolutions to discover important temporal patterns, the proposed architecture employs residual connections as inspired by ResNet [23] and the application of global average pooling (GAP) before classification.

The input data is first convolved with two different-sized kernels, 64 (1x3) and 64 (1x1) to capture its different snapshots. Before applying the 1x3 kernel, the input was convolved with a 1x1 kernel to reduce the overall complexity of the model. Another motivation for the second 1x1 kernel is to make the output feature maps equal in shape. The feature maps yielded by the application of 1x1 and 1x3 are then added followed by Relu activation. After the activation, a dropout with a probability of 50% is applied followed by 1x1 and 1x3 convolutions. A residual connection combines the feature map generated after the dropout with the feature map generated through convolution. To preserve the spatial dimensions, the

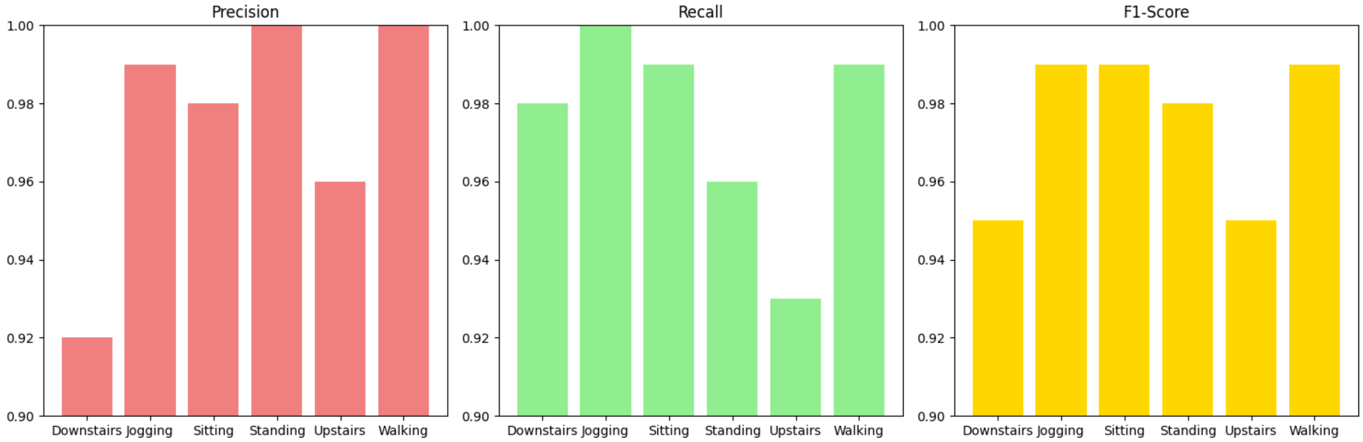

Fig. 3. Performance Scores for each class of WISDM 2011 Dataset: The average Recall, Precision and F1-Scores are 97.5% each

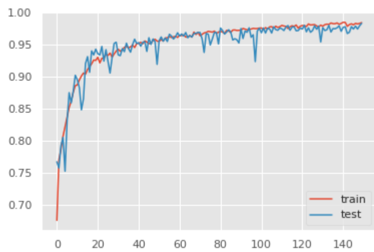

Fig. 4. The accuracies plot for test and training data, the training and testing curves show a good fit, ruling out over-fitting.

padding was kept "same" in our implementation. The output generated after addition is passed to the Relu activation function, which is followed by the global average pooling layer. Finally, a SoftMax layer is employed for class-wise prediction.

The addition of residual connections in our proposed network provides it the capability of feature empowerment and facilitate training. The addition of dropout makes the network more robust by providing the effect of regularization. The best value for the dropout parameter was sought empirically in this work, where a value of 0.5 was proven to yield the best results in our settings.

B. Motivation For Global Average Pooling (GAP) Layer

The motivation behind the adoption of the GAP layer instead of fully connected layers was to keep the total number of parameters to a minimum. To demonstrate the effectiveness of the GAP layer in comparison to its counterpart, we present their comparison in Table I. In Table I, the proposed network is called HARResConnect-Net-GAP when it utilizes the GAP, whereas it is termed HARResConnect-Net-Complex for the case of the fully connected layer. For HARResConnect-Net-Complex, a fully connected layer with 1024 neurons was employed.

It is evident from Table I, that the use of the GAP layer considerably improved the performance of the model while significantly reducing its complexity. Specifically, performance improvements of 4.62% in terms of accuracy along with reductions of > 200 times parameters in terms of complexity were obtained by employing GAP.

C. Model Complexity

Many applications of HAR require low latency at the inference stage making the model preferable for less compute-intensive edge devices including smart-watches and embedded platforms. Therefore, the main goal of this work was to design a model that is less complex than the existing ones while simultaneously having comparable performance. The table II shows the layer-wise trainable parameters along with the dataset used to compute them.

D. Datasets Used

The dataset employed in this work is thoughtfully divided into three portions: 70% for training, 15% for validation, and 15% for testing. The details of the individual dataset are presented in subsequent sections.

This particular dataset, gathered under controlled laboratory conditions, is notable for its absence of missing values, thereby obviating the need for extensive preprocessing. However, it is important to note that the dataset exhibits a pronounced imbalance. The "Walking" class, for instance, is significantly dominant, comprising 424,397 entries, translating to 38.6% of the entire dataset. In a similar vein, the "Jogging" class constitutes 342,176 entries, representing 31.15% of the dataset. The remaining classes exhibit their own proportions, with "Upstairs" accounting for 122,869 entries (11.18%), "Downstairs" encompassing 100,427 entries (9.14%), "Sitting" maintaining 59,939 entries (5.45%), and "Standing" featuring 48,395 entries (4.40%) within the dataset. This class distribution underscores the dataset's inherent imbalance and necessitates specialized consideration in the analysis.

IV. Results

All experiments were conducted on Google's Colaboratory which GPU set as a runtime instance. The number of epochs

TABLE I

BENEFITS OF GLOBAL AVERAGE POOLING: COMPARING THE TRAINABLE PARAMETERS AND PERFORMANCE METRICS OF TWO MODELS, HARRESCONNECT-NET-GAP AND HARRESCONNECT-NET-COMPLEX. THE USE OF GLOBAL AVERAGE POOLING (GAP) IN HARRESCONNECT-NET-GAP RESULTS IN A MUCH LOWER NUMBER OF TRAINABLE PARAMETERS (38,342), WHICH HELPS TO IMPROVE MODEL EFFICIENCY. FURTHERMORE, HARRESCONNECT-NET-GAP OUTPERFORMS HARRESCONNECT-NET-COMPLEX IN TERMS OF AVERAGE ACCURACY, F1 SCORE, RECALL, AND PRECISION PERCENTAGES. THE ADVANTAGES OF ADOPTING GLOBAL AVERAGE POOLING MAY BE SEEN IN THE SIMPLIFIED MODEL DESIGN AND IMPROVED PERFORMANCE MEASURES.

Model	**Trainable Parameters**	**Avg. Accuracy %**	**Avg. F1 %**	**Avg. Recall %**	**Avg. Precision %**
HARResConnect-Net-GAP	38342	98.32	97.5	97.5	97.5
HARResConnect-Net-Complex	8433734	93.7	92.5	90.66	91.66

TABLE II

WISDM 2011 DATASET LAYERWISE PARAMETERS: THE LAYERWISE PARAMETERS OF A NEURAL NETWORK TRAINED ON THE WISDM 2011 DATASET. EACH ROW REFERS TO A SPECIFIC LAYER AND CONTAINS INFORMATION ABOUT THE TRAINABLE PARAMETERS CONNECTED WITH IT. INPUT, CONV1D WITH VARIOUS KERNEL SIZES, ADDITION, ACTIVATION (RELU), GLOBAL AVERAGE POOLING, AND THE FINAL PREDICTION LAYER (DENSE) ARE AMONG THE LAYERS. THE "TOTAL TRAINABLE PARAMETERS" COLUMN AT THE BOTTOM TOTALS 38,342, OFFERING A QUICK INSIGHT OF THE MODEL'S COMPLEXITY AND RESOURCE REQUIREMENTS.

Layer	**Trainable Parameters**
Input Layer	0
Conv1D (1x1)	640
Conv1D (1x3)	12352
Conv1D (1x1)	256
Add	0
Activation (Relu)	0
Conv1D (1x1)	12352
Conv1D (1x3)	12352
Add	0
Activation (Relu)	0
Global Average Pooling	0
Predictions (Dense)	390
Total Trainable Parameters	**38342**

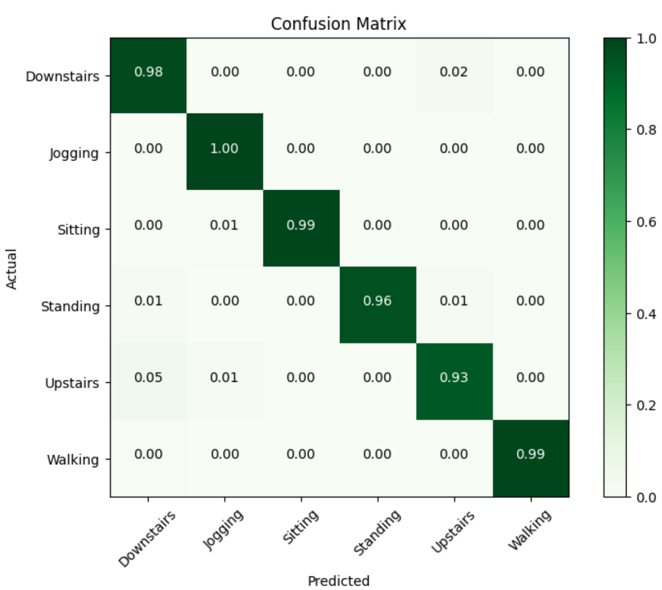

Fig. 5. Confusion matrix showing the classification performance of a model on activity recognition using the WISDM (Wireless Sensor Data Mining) dataset. The matrix illustrates the distribution of actual and predicted activity labels including Downstairs, Jogging, Sitting, Standing, Upstairs, and Walking, with corresponding accuracies highlighted in shades of green.

used for the final training process was 150 and batch size was set to 120. We have selected the number of epochs based on empirical analysis. Experiments were also conducted using early stopping. The in-depth analysis of training and validation loss and accuracy helped us to select the epoch count. Figure shows the accuracy plot for testing and training. It can be seen from it that the model does not overfit. We experimented with different optimizers including "NAdam", "Adam", "Momentum and "RMSprob" in this work and found "Adam" to be the best-suited in our settings.

The best accuracy achieved for WISDM 2011 dataset was 98.32%. Table 5 shows the confusion matrix. The most confused classes are Downstairs and Upstairs, which have a confusion percentage of only 4.8%. The reason for the confusion is that the movement of the body while doing these two instances is nearly identical.The top performing instance is Jogging, which has a classification accuracy of 100%. The reason is that the activity is different from other class with maximum inertial movement involved in it. Which makes it easy to classify. We reported F1, recall, and precision since the dataset was uneven. The complete report is depicted in Figure 3. The average F1, recall, and precision values are all 98%.

TABLE III

PERFORMANCE COMPARISON OF VARIOUS DEEP LEARNING MODELS FOR HUMAN ACTIVITY RECOGNITION (HAR)

Model	**Year**	**Accuracy (%)**
CNN [24]	2018	93.32
CNN [25]	2019	96.40
CNN [26]	2020	95.26
LSTM-CNN [27]	2020	95.85
CNN [2]	2020	94.65
BiGRU-CNN [28]	2022	97.20
CNN [29]	2022	89.60
CNN [30]	2023	94.00
HARResConnect-Net	2023	**98.32**

V. CONCLUSIONS

The provided table III displays a variety of deep learning models used for Human Activity Recognition (HAR) and their corresponding accuracies. The HARResConnect-Net model, in particular, stands out with an accuracy of 98.32%. When compared to the previously established models listed in the table, the HARResConnect-Net outperforms them significantly. The closest competitor, for example, was the BiGRU-CNN model from the paper cited as citeimran2022ultanet, which achieved an accuracy of 97.2% in 2022. This model is outperformed by the HARResConnect-Net by a significant margin of 1.12%.

Furthermore, when the LSTM-CNN model [27] was reviewed in 2020, it achieved an accuracy of 95.85%, indicating a significant improvement in accuracy by HARResConnect-Net, displaying a 2.47% increase. This consistent trend is seen across the other models mentioned in the table, establishing the HARResConnect-Net as a leading HAR model due to its exceptional accuracy.

The evolution of HAR models over time, as shown in the table, demonstrates the remarkable progress that culminated in the HARResConnect-Net's performance. The HARResConnect-Net, developed in 2023, not only has the highest accuracy among all models listed, but it also represents a significant advancement in accurately recognising human activities. This higher accuracy implies that the HARResConnect-Net model may provide more reliable and precise identification of various human movements and actions, establishing itself as a more effective and superior model for Human Activity Recognition when compared to the studies previously presented in the table.

References

[1] J. Lu, X. Zheng, M. Sheng, J. Jin, and S. Yu, "Efficient human activity recognition using a single wearable sensor," *IEEE Internet of Things Journal*, vol. 7, no. 11, pp. 11 137–11 146, 2020.

[2] K. Mehmood, H. A. Imran, and U. Latif, "Hardensenet: A 1d densenet inspired convolutional neural network for human activity recognition with inertial sensors," in *2020 IEEE 23rd International Multitopic Conference (INMIC)*. IEEE, 2020, pp. 1–6.

[3] S.-M. Lee, S. M. Yoon, and H. Cho, "Human activity recognition from accelerometer data using convolutional neural network," in *2017 ieee international conference on big data and smart computing (bigcomp)*. IEEE, 2017, pp. 131–134.

[4] Z. F. Syed, P. Aggarwal, X. Niu, and N. El-Sheimy, "Civilian vehicle navigation: Required alignment of the inertial sensors for acceptable navigation accuracies," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 6, pp. 3402–3412, 2008.

[5] Y. Liu, M. Dashti, M. A. Abd Rahman, and J. Zhang, "Indoor localization using smartphone inertial sensors," in *2014 11th Workshop on Positioning, Navigation and Communication (WPNC)*. IEEE, 2014, pp. 1–6.

[6] J. W. Lockhart, G. M. Weiss, J. C. Xue, S. T. Gallagher, A. B. Grosner, and T. T. Pulickal, "Design considerations for the wisdm smart phone-based sensor mining architecture," in *Proceedings of the Fifth International Workshop on Knowledge Discovery from Sensor Data*, 2011, pp. 25–33.

[7] G. M. Weiss, K. Yoneda, and T. Hayajneh, "Smartphone and smartwatch-based biometrics using activities of daily living," *IEEE Access*, vol. 7, pp. 133 190–133 202, 2019.

[8] D. Anguita, A. Ghio, L. Oneto, X. Parra, J. L. Reyes-Ortiz *et al.*, "A public domain dataset for human activity recognition using smartphones." in *Esann*, vol. 3, 2013, p. 3.

[9] A. Reiss and D. Stricker, "Introducing a new benchmarked dataset for activity monitoring," in *2012 16th international symposium on wearable computers*. IEEE, 2012, pp. 108–109.

[10] A. Stisen, H. Blunck, S. Bhattacharya, T. S. Prentow, M. B. Kjærgaard, A. Dey, T. Sonne, and M. M. Jensen, "Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition," in *Proceedings of the 13th ACM conference on embedded networked sensor systems*, 2015, pp. 127–140.

[11] 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, 2018.

[12] Q. Riaz, M. Z. U. H. Hashmi, M. A. Hashmi, M. Shahzad, H. Errami, and A. Weber, "Move your body: Age estimation based on chest movement during normal walk," *IEEE Access*, vol. 7, pp. 28 510–28 524, 2019.

[13] Q. Riaz, A. Vögele, B. Krüger, and A. Weber, "One small step for a man: Estimation of gender, age and height from recordings of one step by a single inertial sensor," *Sensors*, vol. 15, no. 12, pp. 31 999–32 019, 2015.

[14] I. Gohar, Q. Riaz, M. Shahzad, M. Zeeshan Ul Hasnain Hashmi, H. Tahir, and M. Ehsan Ul Haq, "Person re-identification using deep modeling of temporally correlated inertial motion patterns," *Sensors*, vol. 20, no. 3, p. 949, 2020.

[15] M. A. Hashmi, Q. Riaz, M. Zeeshan, M. Shahzad, and M. M. Fraz, "Motion reveal emotions: Identifying emotions from human walk using chest mounted smartphone," *IEEE Sensors Journal*, vol. 20, no. 22, pp. 13 511–13 522, 2020.

[16] Z. Zhang, Y. Song, L. Cui, X. Liu, and T. Zhu, "Emotion recognition based on customized smart bracelet with built-in accelerometer," *PeerJ*, vol. 4, p. e2258, 2016.

[17] M. Z. U. H. Hashmi, Q. Riaz, M. Hussain, and M. Shahzad, "What lies beneath one's feet? terrain classification using inertial data of human walk," *Applied Sciences*, vol. 9, no. 15, p. 3099, 2019.

[18] V. Bianchi, M. Bassoli, G. Lombardo, P. Fornacciari, M. Mordonini, and I. De Munari, "Iot wearable sensor and deep learning: An integrated approach for personalized human activity recognition in a smart home environment," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8553–8562, 2019.

[19] D. Tao, Y. Wen, and R. Hong, "Multicolumn bidirectional long short-term memory for mobile devices-based human activity recognition," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1124–1134, 2016.

[20] O. Barut, L. Zhou, and Y. Luo, "Multitask lstm model for human activity recognition and intensity estimation using wearable sensor data," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8760–8768, 2020.

[21] A. Shrestha, H. Li, J. Le Kernec, and F. Fioranelli, "Continuous human activity classification from fmcw radar with bi-lstm networks," *IEEE Sensors Journal*, vol. 20, no. 22, pp. 13 607–13 619, 2020.

[22] O. Banos, J.-M. Galvez, M. Damas, H. Pomares, and I. Rojas, "Window size impact in human activity recognition," *Sensors*, vol. 14, no. 4, pp. 6474–6499, 2014.

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[24] A. Ignatov, "Real-time human activity recognition from accelerometer data using convolutional neural networks," *Applied Soft Computing*, vol. 62, pp. 915–922, 2018.

[25] H. Zhang, Z. Xiao, J. Wang, F. Li, and E. Szczerbicki, "A novel iot-perceptive human activity recognition (har) approach using multihead convolutional attention," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1072–1080, 2019.

[26] H. A. Imran and U. Latif, "Hharnet: Taking inspiration from inception and dense networks for human activity recognition using inertial sensors," in *2020 IEEE 17th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET)*. IEEE, 2020, pp. 24–27.

[27] K. Xia, J. Huang, and H. Wang, "Lstm-cnn architecture for human activity recognition," *IEEE Access*, vol. 8, pp. 56 855–56 866, 2020.

[28] H. A. Imran, "Ultanet: A antithesis neural network for recognizing human activity using inertial sensors signals," *IEEE Sensors Letters*, 2022.

[29] H. A. Imran, "Khail-net: A shallow convolutional neural network for recognizing sports activities using wearable inertial sensors," *IEEE Sensors Letters*, vol. 6, no. 9, pp. 1–4, 2022.

[30] H. A. Imran, A. A. Ikram, S. Wazir, and K. Hamza, "Edgeharnet: An edge-friendly shallow convolutional neural network for recognizing human activities using embedded inertial sensors of smart-wearables," in *2023 International Conference on Communication, Computing and Digital Systems (C-CODE)*, 2023, pp. 1–6.