

ECG-NETS – A novel integration of capsule networks and extreme gated recurrent neural network for IoT based human activity recognition

S. Arokiaraj^{a,*} and N. Viswanathan^b

^a*Department of Electronics and Communication Engineering, Tagore Institute of Engineering and Technology, Salem, Tamil Nadu, India*

^b*Department of Electronics and Communication Engineering, Mahendra Engineering College (Autonomous) Namakkal, Tamil Nadu, India*

Abstract. Human Activity Recognition (HAR) has reached its new dimension with the support of Internet of Things (IoT) and Artificial Intelligence (AI). To observe human activities, motion sensors like accelerometer or gyroscope can be integrated with microcontrollers to collect all the inputs and send to the cloud with the help of IoT transceivers. These inputs give the characteristics such as, angular velocity of movements, acceleration and apply them for an effective HAR. But reaching high recognition rate with less complicated computational overhead still represents a problem in the research. To solve this aforementioned issue, this work proposes a novel ensembling of Capsule Networks (CN) and modified Gated Recurrent Units (MGRU) with Extreme Learning Machine (ELM) for an effective HAR classification based on data collected using IoT systems called Ensemble Capsule Gated (ECG)-Networks (NETS). The proposed system uses Capsule networks for spatio-feature extraction and modified (Gated Recurrent Unit) GRU for temporal feature extraction. The powerful feed forward training networks are then employed to train these features for human activity recognition. The proposed model is validated on real time IoT data and WISDM datasets. Experimental results demonstrates that proposed model achieves better results comparatively higher than existing (Deep Learning) DL models.

Keywords: Artificial intelligence, capsule networks, human activity recognition, internet of things, gated recurrent units and spatio-feature extraction

1. Introduction

In recent days, there is a huge advancement in human activity recognition (HAR) systems using interconnected sensing and intelligent technologies such as artificial intelligence (AI), Internet of Things (IoT), and sensors. Wearable IoT devices with sen-

sors plays very vital role to observe the movements of human body and also can be aided for its recognition. For this kind of research, Inertial Measurement units (IMU) are integrated with wearable IoT devices to identify human activities which enables development of different cognitive applications in areas such as healthcare [1], fitness application [2] and security applications [3]. The integration of IoT in human activity recognition has been used for an effective human activity collection which are implemented for an effective recognition system.

*Corresponding author. S. Arokiaraj, Department of Electronics and Communication Engineering, Tagore Institute of Engineering and Technology, Salem, Tamil Nadu, India. E-mail: nithiarokia@gmail.com.

Recently IoT merged with machine and DL algorithms for an automatic human activity recognition. Several previous studies investigated deployment of HAR using traditional machine algorithms such as Support Vector Machines (SVM) [4], Artificial Neural Networks (ANN) [5], Random Forest (RF) [6], Convolutional Neural Networks (CNN) [7], Recurrent neural Networks (RNN) [8] and Long Short-Term Memory (LSTM) [9]. Since the HAR consists of simple and complex activities, employing the DL algorithms for task of classifying the human activities are considered to be more suitable than the machine learning algorithms. In recent times, hybrid DL algorithms such as CNN-LSTM [10], CNN-GRU [11], and transfer learning techniques are employed to handle the complex activities [12]. But these algorithms requirement, its improvisation to achieve better HAR with less computational complexity.

The idea behind this research is to develop the cognitive and automatic system with the high recognition rate and low computational complexity. The research integrates the Capsule Networks (CN) with the Modified Gated Recurrent Units (MGRU) to extract highly efficient spatio-temporal features and trained with powerful feed-forward networks. The research contributions are listed as follows.

The paper discusses hybrid DL model which consist of Convolutional Capsule Networks and Modified Gated Recurrent Units (MGRU) for an effective extraction of spatio-temporal features from the raw IoT data.

- The paper leverages the Extreme Feed- Forward Networks for a high-speed training with the less complexity for deep-learned features.
- The real time IoT wearable test beds are proposed for an effective data collection which represents various human activities.
- The evaluation of performance measures are achieved and compared with existing HAR systems based on DL.

The paper structure is presented as follows. Section II presents the papers from the different authors. 2) The working mechanism of the proposed framework, dataset descriptions, data preprocessing and hybrid DL model is explained in Section III. The implementation is described in section IV and the experimental results, comparative analysis and discussions are provided in the Section V. The conclusive remarks and future enhancement are summarized in Section-VI.

2. Related works

Abdullah et al. developed CNN for HAR utilizing their strolling styles which is recognized at lower appendages. This framework adopted Butterworth filter for NN training and Levenberg–Marquardt algorithm is utilized for the extraction of real time samples. This framework achieved better performance in terms of performance. However, this architecture has a drawback in that it is inappropriate for dynamic walk patterns [13].

Aybuke et al. tested HuGaDB datasets using Random Forest (RF), Naive Bayes (NB) and IB1 classifiers for the HAR system. This dataset comprises the data of standing, sitting, running and walking. These data are collected by utilizing inertial sensors like accelerometers and gyroscopes. While doing the comparative analysis, it was concluded that RF outperformed NB and IB1 classifier in regards of classification accuracy and setup time [14]. Jucheol et al. presented a multi-model walk ID classifier in view of RNN joined with a SVM extractor [15].

Jiang and Yin created time frequency spectral images form the time serial image by incorporating Short-time Discrete Fourier Transform (STDFT) [16] for identifying basic activities such as waling and standing using CNN method. Integration of CNN and STDFT made perfect HAR system. Main advantage of this framework is it achieved classification accuracy of 95.25 in 25 atomic hand activity classification with the help of 12 participants [17]. The otherworldly properties can be utilized for wearable sensor movement acknowledgment, yet in addition for action acknowledgment without utilization of a gadget. Ha and Choi [18] presented a learning methodology explicit transient property for a new version of CNN for HAR. This framework provided better result in terms of classification accuracy for HAR.

Shen et al. [19] utilized gated CNN to perceive ordinary exercises from sound signals more precise than CNN. Long et al. made a two-stream CNN structure with the help of residual blocks that can deal with a few time scales. Guan and Plötz [20] developed deep LSTM network to test various HAR benchmark datasets to outstand other frameworks. Yao et al. [21] utilized RNN integrated GRU framework instead of LSTM for HAR system. But in few cases, LSTM framework outperformed other versions of RNN based on accuracy classification. Wang et al. [23] utilized sound features for the HAR system. In this framework CNN is integrated with LSTM to cap-

ture various gesture movements of human. Xu et al. [24] utilized contemporary Inception CNN structure, while GRUs are utilized for effective global temporal portrayals.

Yuki et al. [25] utilized a double stream ConvLSTM organization responsible for different time length. Alazrai et al. [26] suggested MLPs are utilized to generate a base classifier for every physical methodology, and selected at classifier level to consolidate all classifiers. Accordingly, the variety of various modalities is held, which is significant for defeating over-fitting hardships and upgrading by and large speculation limit. Xia et al. [27], notwithstanding the conventional classifier group, zeroed in on the fall discovery issue and presented a troupe of the remaking mistake from the auto encoder of sensor methodology.

3. Proposed methodology

3.1. Overview of proposed system architecture

The proposed framework includes four major steps, namely IoT based Data Collection Unit (DCU), Data pre-processing, spatio-temporal feature extraction using Capsule and Modified GRU units and HAR classification mechanism. Figure 1 shows proposed architecture for ECG-NETS that are used for an effective human activity recognition system.

3.1.1. Data collection mechanism

Nearly 55 persons with body weight ranging from 30 kg to 65 kg are considered. The battery -operated IoT devices are used for collecting the body activities. Table 1 presents the hardware with its specifications and applications for the experimentation. To collect daily human activities, ADXL435 3-axis accelerometers and BMG250 3 axis gyroscopes are interfaced with MCP3008 (10-bit ADC) connected Node MCU. In order to collect data, Micro-Python programming is deployed in hardware and collected data is transmitted to the cloud. 3.3 V Li-On batteries are utilized to power up the board. Nearly 8,20,567 data are collected from above subjects for 2 months conducting the various human activities such as eating, folding the clothes, clapping and even brushing the teeth. Each subject is permitted to play out the 18 distinct exercises for 2 minutes at 15 Hz which determines that 25 information focuses are recorded for every second of activity being performed. This sensor information gathered from IoT gadgets are put away in

the Cloud by making the proper information outlines. The properties data in the cloud is depicted in Table 2. Once the data are stored in the cloud, these data are downloaded for further process. Figures 2 and 3 show sample data distribution for hand activities using IoT test beds over the time period.

3.1.2. Data pre-processing

The next step is Data-preprocessing which is used to organize data suitable for training and testing the module. Since dataset is organized in flat databases, these datasets are converted into suitable data frames using the Pandas library in Python. The missing values and zero row values were removed from the datasets. The noise values are removed from the datasets by comparing the original values of sensors. The datasets collected are evenly distributed to avoid the class imbalance problem. The information is sectioned utilizing a sliding window of 05 seconds and half cross-over by utilizing a step size of 100. This progression is essential with the goal that information can be changed into a fitting configuration of time steps and highlights for input into DL models. The first dataset will be divided into training, testing, and validation (70 : 20 : 10) as the objective is to generate models that can manage data from clients. For training and testing, the DL models utilized in this work are adjusting their hyper-boundaries. Finally, testing is performed on test datasets to acquire results.

3.2. Proposed models

The gathered datasets embody the estimations of the amount as pieces of information over the long run. The collected data points over the time determines the behavior of the data. These time series datasets are incorporated to train the model and to achieve better classification, hybrid DL algorithms ensemble with machine learning algorithms to extract features and applied for segmentation.

3.2.1. ECG-NETS- Hybrid models

Hybrid Models used in this research are an amalgamation of deep and machine learning models [28]. The proposed model integrates the combination of Capsule networks and Modified GRU models to extract the spatio-temporal features processed by “Single Feed forward Extreme Learning Machines” (ELM) [29] for the further classification. The preliminary overview of each learning model is discussed in the following section.

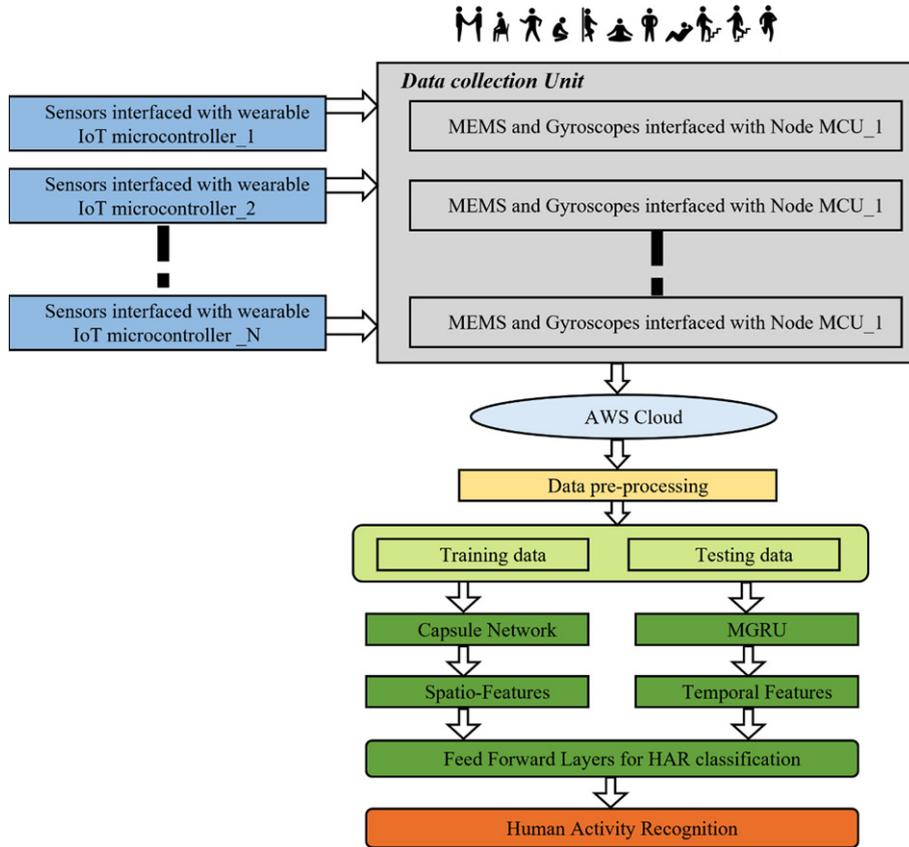


Fig. 1. Proposed architecture for ECG-NETS.

Table 1
Specification of the hardware used for an effective data collection unit

S. No	Hardware Used	Specifications	Application
01	Node MCU	8-bit CPU with ESP8266 WIFI transceivers	Considered as the main CPU for interfacing with ADC and other sensors
02	MCP3008	10-bit ADC with 8-Input Channels	Used for the interfacing the sensors with Node MCU using Serial Peripheral Interfaces (SPI)
03	ADXL435	Three Axis Accelerometers	Measures the body activities of the subjects.
04	BMG250	Three Axis Gyroscopes are used	
05	Power Supply	3.3 V Li-ON batteries	Power Supply Unit for the Wearable IoT Boards.

3.2.2. Capsule networks

Recently, Capsule networks [30] are gaining the edge of popularity over the Convolutional neural networks in extracting the spatial features from the group of datasets. Conventional CNN has three layers such as “Convolutional layers (CL), Pooling layers (PL) and Fully Connected layers (FCL)”. The features are extracted by the early two layers which aided by the FCL for the further classification. In this process, CNN extracts all the features but may leads

to non-consideration of spatial information as the datasets increases due to the nature of pooling layers in CNN.

For improved component spatial extraction, the pooling layers of CNN are replaced with capsule networks. Furthermore, Capsule Network is also proposed as a solution to CNN problem. The clusters of neurons known as capsules are responsible for encoding both spatial information and the likelihood that an item is always accessible based on factors including,

Table 2
Attribute description of the data stored in the cloud

S. No	Attributes obtained	Description
01	Subject ID	Represents the details of Subjects
02	Sensors X1, Y1, Z1	Sensors obtained from the accelerometers
03	Sensors X2, Y2, Z2	Sensors obtained from Gyroscopes
04	Activity label -U	U- values ranges from 0 to 17 which represents the 18 activities

- Instantiation boundaries of entities
- Probability of Existence

Three layers make up the capsule network: a low capsule layer, a high container layer, and a classification layer. Global boundary sharing is used to reduce the occurrence of errors and an upgraded dynamic steering computation is used to iteratively update the boundaries.

The increase of the information vectors' framework with weighted lattice is determined in order to encode fundamental spatial relationship between low and substantial level convolutional highlights in image is given by,

$$Y(i, j) = W_i(i, j) * S_j \tag{1}$$

Where $Y(i, j)$ is information vector with weight lattice, $W_i(i, j)$ is the weighted coefficient matrix

that is learned in the backward process, S_j is parent capsule. The current capsules are calculated by adding input weight vectors and forwarded its output to higher level capsule as expressed in Equation (2).

$$J = \sum j(i, j) * (j) \tag{2}$$

where, i and j are input and output weight vectors respectively.

From Equation (3), non-linearity applies the squash function which reduces a vector's length to a maximum of one and a minimum of zero while preserving its orientation.

$$j = \text{squash}[S(j)] \tag{3}$$

where, S (is the squash function).

3.2.3. GRU model based feature extraction

The GRU model plays vital role in extraction of temporal features, which gathers information received from IoT cloud. Figure 4 illustrates GRU structure utilized in this framework. GRU network consists of two gates and which are faster than LSTM [32] and RNN models [33]. Equations (4-7) are used to describe the characteristics of GRU,

$$z_t = \sigma(w_z \cdot [h_{t-1}, x_t]) \tag{4}$$

$$r_t = \sigma(w_r \cdot [h_{t-1}, x_t]) \tag{5}$$

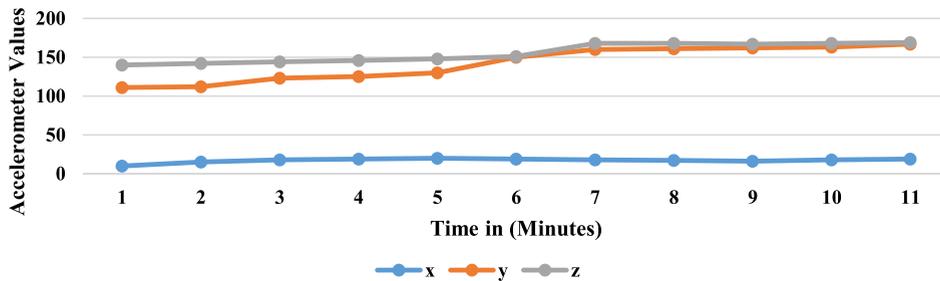


Fig. 2. Sample data collected for hand activities (accelerometers) using IoT test beds.

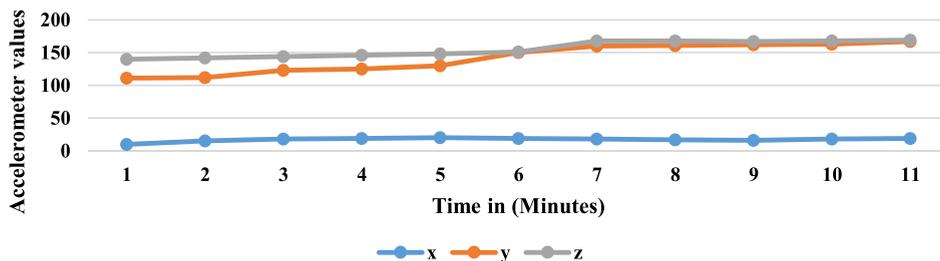


Fig. 3. Sample data collected for hand activities (gyroscopes) using IoT test beds.

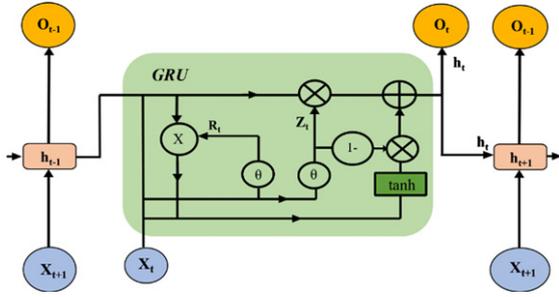


Fig. 4. GRU model for ECG-NETS architecture.

$$h_t = \tanh(r_t * [h_{t-1}, x_t]) \tag{6}$$

$$(1 - z_t) \otimes h_{t-1} + z_t \otimes h_t \tag{7}$$

where z_t means update gate, x_t means current input vector, h_{t-1} , means value measured form previous adjacent layer, w_z means update gate’s learning weight matrix, r_t means reset gate, \tanh ranges from $(-1$ to $1)$, h_t means the present output of module.

3.3. MGRU: Capsule – GRU feature extraction

For the effective spatial information, capsule networks extract the different information and obtain the relationship between the features. Figure 5 shows the hybrid feature extractor using capsule GRU networks.

The result weights are determined utilizing equation (2) and it is transferred to high case region while squash work holds the first course of the vector by packing the length to $(0, 1)$ utilizing equation (3). In the following stage, proposed model consolidates spot item between comparative cases and results, utilizing advanced dynamic steering. This cycle iteratively refreshes the loads of the organizations to frame the component maps. At long last the component maps from the case network are gone through the dropout layer to forestall the over-fitting and leveled to a solitary layered highlights vectors. These component vectors go about as the contribution to the proposed Modified GRU. The decision of utilizing a GRU layer is that it has a lesser number of gates when contrasted with LSTM and hence has similarly less boundaries than LSTM and RNN. In this exploration article, Modified GRU is utilized in which each hidden node is figured utilizing just the inclination, accordingly decreasing the absolute number of boundaries. The compositional system for the proposed model is displayed in Fig. 4.

The proposed Hybrid Feature Extractor Structure using Capsule -GRU networks is depicted in Fig. 5. Now these feature vectors act as the input to the Modified GRU to extract the further temporal features followed by dropout and dense extreme learning classifiers. The working mechanism of the proposed Capsule-GRU feature extraction in the proposed network is presented in Algorithm-1.

Step No.	Algorithm-1 //Procedure for the Capsule-GRU networks
1	Inputs : Preprocessed Data from the Cloud $I=\{X_1, X_2, X_3 \dots \dots \dots X_n\}$
2	Outputs : feature maps F
3	while $n = 1$ to Max_iteration
4	$H = \text{CNN}(X_1, X_2, X_3, X_4 \dots \dots X_n)$
5	$Y = \text{Capsule}(H)$
6	$Z = \text{flatten}(Y)$
7	$F1 = \text{ModifiedGRU}(Z)$
8	$F = \text{flatten}(F1)$
9	End

3.4. ECG-NETS- classification layers

The final feature maps are then feed to the Classification layers to categorize the 18 different human activities. The proposed model uses the Extreme learning machine based feed forward networks. A type of neural network called an ELM uses a single hidden layer and operates based on auto-tuning properties.

When compared to other learning models including ELM performs faster and more effective with less computing overhead when compared to other learning models like SVM, k-NN and RF. The input features maps of the ELM are denoted using Equation (8)

$$X = G(F, P) \tag{8}$$

where, G refers to GRU function; P is GRU temporal features; F is CNN spatial features and X is fused spatio-temporal GRU features obtained from CNN and GRU layers.

The output of ELM function is represented by Equation (9)

$$Y(n) = X(n) \beta = X(n) X^T \left(\frac{1}{C} X X^T \right)^{-1} O \tag{9}$$

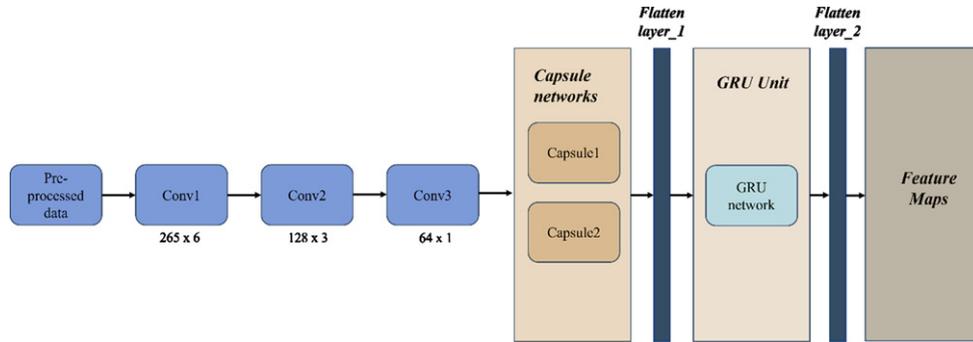


Fig. 5. Hybrid feature extractor structure using capsule GRU Networks.

The complete training of ELM is assumed as Equation (10)

$$S = \alpha \left(\sum_{n=1}^N (Y(n), B(n), W(n)) \right) \quad (10)$$

Where B and W are weights and bias factors of the network with activation function, β is temporal matrix based on inverse theorem of Moore–Penrose algorithm denoted by X^T , X (n) is input feature maps and C is constant.

4. Implementation description

For evaluation, experimental tests are performed using Google Co-lab, Python 3.8, Keras 2.4.3, Numpy 1.19, Pandas 1.12 and Tensor-Flow 2.3.0 with 14.76 GB RAM and Tesla T4 GPU. To evaluate the proposed model, two datasets such as IoT based Datasets and WISDM datasets which consist of the smart phone and smart watch datasets. Table 3 presents the datasets descriptions used for the evaluation.

5. Model evaluation

Table 4 presents the performance metrics utilized for preparing the proposed network. The performance

Table 3

Dataset descriptions used for the evaluation and experimentation

Dataset Details	No of Datasets Collected	Data Splitting (Training: testing: Validation)
Real time Datasets	8,20,567	70 : 20 : 10
WISDM Datasets	8,51,386	70 : 20 : 10

Table 4
Mathematical expressions for the performance metrics' calculation.

Performance Metrics	Mathematical Expression
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Recall	$\frac{TP}{TP+FN}$
Specificity	$\frac{TN}{TN+FP}$
Precision	$\frac{TP}{TP+FP}$

metrics including accuracy, precision, AUC (Area under ROC), recall, specificity and F1-score utilizing different datasets are determined. The early stopping method is utilized in this work to address over-fitting and speculation issues when approval execution shows no improvement for N sequential times [34].

“TP is True Positives, TN is True Negatives, FP is False Positives and FN is False negatives”. The performans measured by utilizing WISDM dataset and ROC and disarray grid is calculated for proposed network model utilizing real-time IoT datasets.

Figures 6 to 8 illustrate the proposed model's Confusion Matrix various datasets. The results cleared that confusion matrix of proposed model has produced 98.45% detection of human activity under different datasets. Figure 8 depicted as confusion matrix for proposed system.

Figure 9 illustrates the proposed model's ROC curves for different HAR datasets. From the ROC curve it is evident that proposed framework shows 0.98 of AUC for IoT data collected, 0.9793 AUC for Smart watch and 0.9823 AUC for smartphone datasets. Figures 10 and 11 illustrate the training and validation curves of proposed ECG-NETS framework in terms of accuracy under IoT datasets and WISDM datasets respectively.

Figures 12 to 16 demonstrate performance comparison of different learning models with proposed

LABEL	Ambulation oriented	Hand oriented eating	Hand oriented general actions	Walking
Ambulation oriented	98.45	1.2	0	0
Hand oriented eating	0	98.56	1.2	0
Hand oriented general actions	1.2	0	98.6	0
Walking	0.45	0.7	0	98.59

Fig. 6. Proposed architecture’s confusion matrix using the real-time IoT datasets.

LABEL	Ambulation oriented	Hand oriented eating	Hand oriented general actions
Ambulation oriented	98.5	1.2	0
Hand oriented eating	0	98.72	1.0
Hand oriented general actions	1.2	1.0	98.70

Fig. 7. Confusion matrix for proposed architecture using the WISDM dataset (smart phone).

LABEL	Ambulation oriented	Hand oriented eating	Hand oriented general actions
Ambulation oriented	98.7	1.3	0
Hand oriented eating	0	98.69	1.3
Hand oriented general actions	1.3	1.0	98.70

Fig. 8. Confusion matrix for proposed architecture using the WISDM dataset (smart watch).

model in detecting HAR under different datasets. CNN+GRU [35], Deep CONV LSTM [36] networks delivered the better order exhibitions. In these two cases, LSTM [38] and inception model [37]

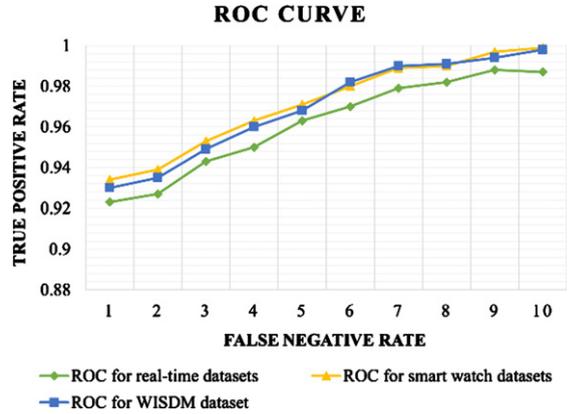


Fig. 9. ROC curves (a) Real-time datasets, (b) smart watch datasets, (c) smart phone datasets (WISDM).

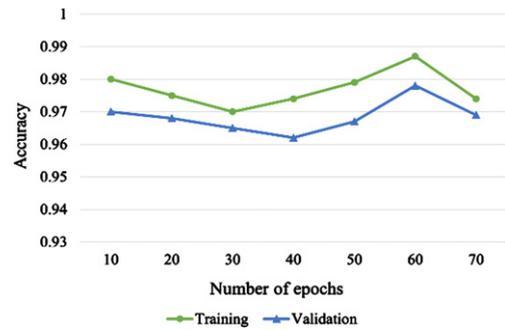


Fig. 10. Validation curves of proposed ECG-NETS model in handling the IoT datasets.

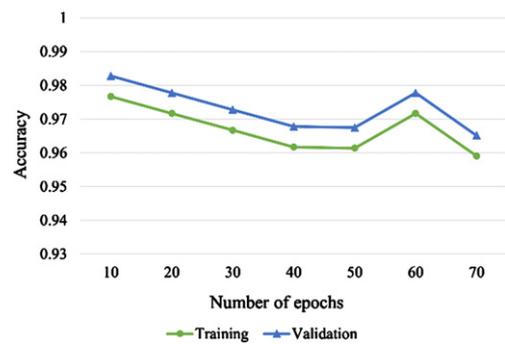


Fig. 11. Validation curves of proposed ECG-NETS model in handling the WISDM datasets.

has delivered minimal execution in identifying the different human exercises. However, ECG-NETS outperformed other existing learning models with quantitative outcomes. The consideration of case networks with the GRU organizations and single feed forward ELM in proposed model effectively affects

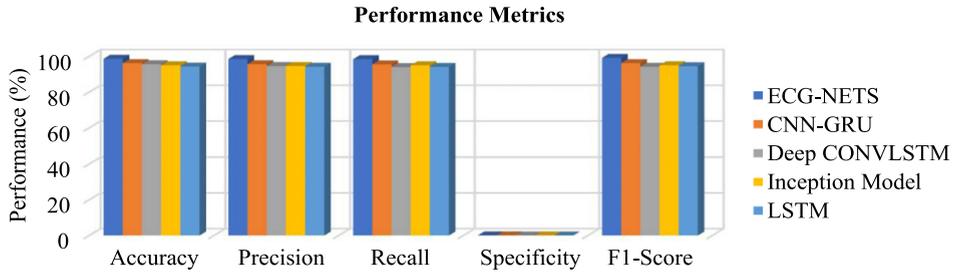


Fig. 12. Performance metrics comparison for exiting methods in detecting ambulation activities (real time datasets).

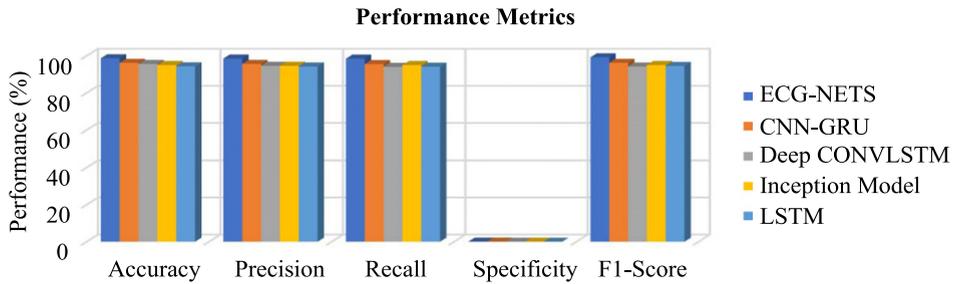


Fig. 13. Performance metrics comparison for exiting methods in detecting hand related activities (real time datasets).

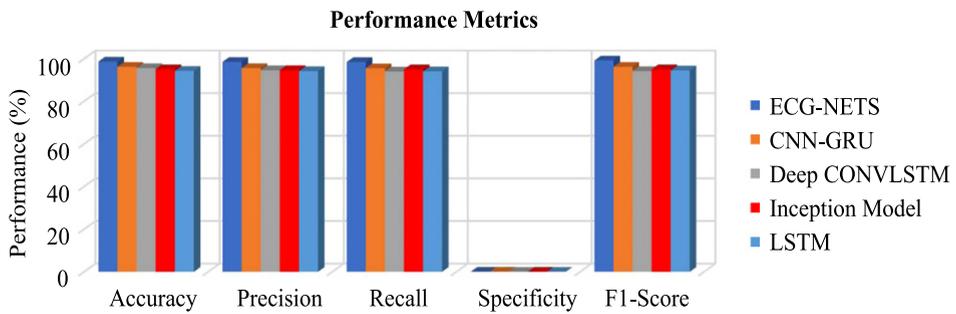


Fig. 14. Performance metrics comparison for exiting methods in detecting hand related eating activities (real time datasets).



Fig. 15. Performance metrics comparison for exiting methods in detecting leg related eating activities (real time datasets).

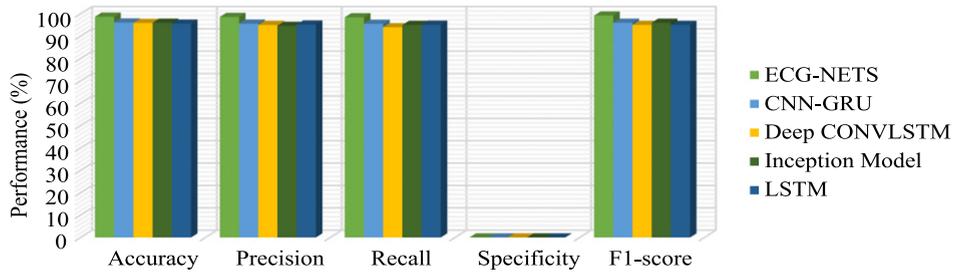


Fig. 16. Performance metrics comparison for exiting methods in detecting various activities using WISDM smart watch and smart phones datasets.

the grouping exhibitions than existing hybrid learning calculations. The similar fashion of performance is observed for the WISDM datasets in which the proposed model has achieved the better performances than the other existing hybrid models.

6. Conclusion and future perspectives

In this work, a novel hybrid DL model ECG-NETS for classification of HAR is proposed by integrating capsule network and GRU unit to accomplish the better extraction of spatio-temporal features. Raw sensor information from IoT test beds and WISDM was utilized in the proposed methodology. Information exchange during preprocessing is performed by utilizing sliding window approach. The ablation analysis is carried out for current models such as Deep ConvLSTM and Inception Time, CNN+GRU and LSTM and different execution measurements are determined. From the results obtained using this proposed framework trained on real-time datasets, it is concluded that the proposed hybrid method offers optimal accuracy over other current models in perceiving complicated human exercises. Therefore, the results demonstrated that hybrid DL models can effectively extract the features. Future directions may also use advanced transformer and models to order the HAR's time series from IoT datasets. The transformer integration will improve the functionality of model so that it can handle number of IoT devices.

References

- [1] N. Ahmed, J.I. Rafiq and M.R. Islam, Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model, *Sensors* **20**(317) (2020), 1–19.
- [2] K. Chen, D. Zhang, L. Yao, B. Guo, Z. Yu and Y. Liu, DL for sensor-based human activity recognition: overview, challenges and opportunities, *ACM Computing Surveys* **54**(4) (2021), 1–40, arXiv:2001.07416.
- [3] K. Chen, L. Yao, D. Zhang, X. Wang, X. Chang and F. Nie, A semi-supervised recurrent convolutional attention model for human activity recognition, *IEEE Transactions on Neural Networks and Learning Systems* **31**(5) (2020), 1747–1756, doi: 10.1109/TNNLS.2019.2927224.
- [4] L. Yao, B. Guo, Z. Yu and Y. Liu, DL for sensor based human activity recognition: overview, challenges, and opportunities, *ACM Computing Survey* **54**(4) (2020), 10.1145/3447744.
- [5] J. Hoey, C.D. Nugent, D.J. Cook and Z. Yu, Sensor-based activity recognition, *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews* **42**(6) (2012), 790–808, 10.1109/TSMCC.2012.2198883.
- [6] A.K. Kushwaha, A.K. Kar and Y.K. Dwivedi, Applications of big data in emerging management disciplines: A literature review using text mining, *International Journal of Information Management Data Insights* **1**(2) (2021), 10.1016/j.jjime.2021.100017.
- [7] M. Organero and R. Blazquez, Time-elastic generative model for acceleration time series in human activity recognition, *Sensors* **17**(2) (2017), 10.3390/s17020319.
- [8] J.A. Nasir, O.S. Khan and I. Varlamis, Fake news detection: A hybrid CNN-RNN based DL approach, *International Journal of Information Management Data Insights* **1**(1) (2021), 1–13, 10.1016/j.jjime.2020.100007.
- [9] H.F. Nweke, Y.W. Teh, M.A. Al-garadi and U.R. Alo, DL algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. *Expert Systems with Applications* **105** (2018), 233–261, 10.1016/j.eswa.2018.03.056.
- [10] R. Ramamurthy and N. Roy, Recent trends in machine learning for human activity recognition a survey, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **8**(4) (2018), 1–11.
- [11] D.C. Ranasinghe, R.L.S. Torres and A. Wickramasinghe, Automated activity recognition and monitoring of elderly using wireless sensors: Research challenges. Proceedings of the 2013 5th IEEE International Workshop on Advances in Sensors and Interfaces, 2018, 224–227.
- [12] S. Ranasinghe, F. Al MacHot and H.C. Mayr, A review on applications of activity recognition systems with regard to performance and evaluation, *International Journal of Distributed Sensor Networks* **12**(8) (2016), 10.1177/1550147716665520.
- [13] A. Alharbi, K. Equbal, S. Ahmad, H.U. Rahman, H. Alyami, Human gait analysis and prediction using the levenberg-marquardt method, *Hindawi*

- Journal of Healthcare Engineering* (2021), 1–11, <https://doi.org/10.1155/2021/5541255>
- [14] A.M. Saleh and T. Hamoud, Analysis and best parameters selection for person recognition based on gait model using CNN algorithm and image augmentation, *Journal of Big Data* **8**(1) (2021), <https://doi.org/10.1186/s40537-020-00387-6>
- [15] J. Moon, N.A. Le, N.H. Minaya, S.I. Choi, Multimodal few-shot learning for gait recognition, *Applied Sciences* **10** (2020), doi:10.3390/app10217619
- [16] W. Jiang and Z. Yin, **Human activity recognition using wearable sensors by deep convolutional neural networks**, In *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 2015, 1307–1310.
- [17] G. Laput and C. Harrison, **Sensing fine-grained hand activity with smartwatches**, In *Proceedings of the CHI Conference on Human Factors in Computing Systems* 2019, 1–13.
- [18] S. Ha and S. Choi, **Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors**, *Proceedings of the International Joint Conference on Neural Networks*, 2016, 381–388.
- [19] Y.H. Shen, K.H. He and W.Q. Zhang, **SAM-GCNN: A Gated convolutional neural network with segment-level attention mechanism for home activity monitoring**, *Proceedings of the IEEE International Symposium on Signal Processing and Information Technology*, 2018, 679–684.
- [20] H. Guo, L. Chen, L. Peng, and G. Chen, **Wearable sensor based multimodal human activity recognition exploiting the diversity of classifier ensemble**, *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, 1112–1123.
- [21] S. Yao, S. Hu, Y. Zhao, A. Zhang, and T. Abdelzaher, **DeepSense: a unified dl framework for time-series mobile sensing data processing**, *Proceedings of the International Conference on World Wide Web, International World Wide Web Conferences Steering Committee*, 2017, 351–360.
- [22] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, LSTM: A search space odyssey, *IEEE Transactions on Neural Networks and Learning Systems* **28**(10) (2016), 2222–2232.
- [23] Y. Wang, J. Shen, and Y. Zheng, **Push the limit of acoustic gesture recognition**, *Proceedings of the IEEE Transactions on Mobile Computing*, 2020, 566–575.
- [24] A. Xu, D. Chai, J. He, X. Zhang, and S. Duan, InnoHAR: a deep neural network for complex human activity recognition, *IEEE Access* **7** (2019), 9893–9902.
- [25] Y. Yuki, J. Nozaki, K. Hiroi, K. Kaji, and N. Kawaguchi, **Activity recognition using dualconvlstm extracting local and global features for sh1 recognition challenge**, *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, 2018, 1643–1651.
- [26] R. Alazrai, M. Hababeh, B.A. Alsaify, M.Z. Ali and M.I. Daoud, An end-to-end deep learning framework for recognizing human-to-human interactions using wi-fi signals, *IEEE Access* **8** (2020), 197695–197710, doi: 10.1109/ACCESS.2020.3034849.
- [27] K. Xia, J. Huang and H. Wang, Lstm-cnn architecture for human activity recognition, *IEEE Access* **8** (2020), 56855–56866, 10.1109/ACCESS.2020.2982225.
- [28] L. Zhang, X. Wu and D. Luo, **Human activity recognition with HMM-DNN model**, *Proceedings of 2015 IEEE 14th International Conference on Cognitive Informatics and Cognitive Computing*, 2015, 192–197, 10.1109/ICCI-CC.2015.7259385.
- [29] G.B. Huang, Q.Y. Zhu, and C.K. Siew, Extreme learning machine: theory and applications, *Neuro-computing* **70**(1) (2006), 489–501.
- [30] A. Pham et al., SensCapsNet: deep neural network for non-obtrusive sensing based human activity recognition, *IEEE Access* **8** (2020), 86934–86946, doi: 10.1109/ACCESS.2020.2991731.
- [31] Bibi, A. Akhuzada, J. Malik, J. Iqbal, A. Musaddiq and S. Kim, A dynamic dl-driven architecture to combat sophisticated android malware, *IEEE Access* **8** (2020), 129600–129612, doi: 10.1109/ACCESS.2020.3009819.
- [32] X. Shu, L. Zhang, Y. Sun and J. Tang, Host-parasite: graph lstm-in-lstm for group activity recognition, *IEEE Transactions on Neural Networks and Learning Systems* **32**(2) (2021), 663–674, doi: 10.1109/TNNLS.2020.2978942.
- [33] A. Wang, S. Huang, J. Qiu, Parallel online sequential extreme learning machine based on MapReduce, *Neuro-computing* **149** (2015), 224–32.
- [34] N. Zhao, H. Gao, X. Wen and H. Li, Combination of convolutional neural network and gated recurrent unit for aspect-based sentiment analysis, *IEEE Access* **9** (2021), 15561–15569.
- [35] A. Pham et al., SensCapsNet: deep neural network for non-obtrusive sensing based human activity recognition, *IEEE Access* **8** (2020), 86934–86946,
- [36] X.Y. Lim, K.B. Gan and N.A. Abd Aziz, Deep ConvLSTM Network with dataset resampling for upper body activity recognition using minimal number of imu sensors, *Applied Sciences* **11**(8), (2021), 1–18, <https://doi.org/10.3390/app11083543>.
- [37] S. Khater, M. Hadhoud and M.B. Fayek, A novel human activity recognition architecture: using residual inception ConvLSTM layer, *Journal of Engineering and Applied Sciences* **69**(45) (2022), 1–16.
- [38] S. Gajjala and B. Chakraborty, **Human activity recognition based on lstm neural network optimized by pso algorithm**, *Proceedings of IEEE 4th International Conference on Knowledge Innovation and Invention*, 2021, 128–133.