J. Electrical Systems 20-9s (2024): 2407-2412

[1] S. Prasanna Bharathi*

[2] G. Chamundeeswari

[3] Patri Upender

Human Activity Recognition using GRU Algorithm

JES Journal of Electrical Systems

Abstract: - The recognition of human activity is a challenging task nowadays, where surveillance cameras are employed for security purposes and monitoring. It is highly appreciated to utilize the best deep learning techniques to achieve higher accuracy than existing methods. In this research, the Gated Recurrent Unit (GRU) technique is employed to identify activities performed by humans. The Human Activity Recognition Trondheim (HARTH) dataset is utilized, which consists of data collected from individuals. The dataset comprises six different classes of daily activities performed by humans: climbing downstairs, upstairs, walking, sitting, running, and standing. The algorithm is implemented against the HARTH dataset to achieve higher accuracy using TensorFlow and the Python framework, and accuracy is calculated. A confusion matrix is also obtained from the conducted research. This research concludes that the GRU algorithm yields a higher accuracy of around 95% in identifying human activities compared to the machine learning algorithms implemented earlier.

Keywords: Deep Learning, Gated Recurrent Unit, HARTH Dataset, RNN, Human action recognition.

I. INTRODUCTION

Identifying human activities is necessary to recognize any suspicious behavior in crowded places and ensure the safety of society. Therefore, research in this area has a significant impact on ensuring safety and security in day-to-day life. From this perspective, deep learning algorithms have become integral in recognizing and assessing human activities in daily life [1]. Human activity recognition spans various fields such as sports, daily routines, remote surveillance, interactive gaming, yoga, fitness monitoring, and healthcare [2-7]. In some areas, wearable accelerometers have been utilized to assess human activity in hospitals and among patients, enabling remote connectivity. Despite the utilization of several machine learning algorithms to address this issue, achieving higher accuracy remains a challenge [8-13]. Although many solutions have been proposed by researchers, real-time capturing of videos and images still falls short in achieving higher accuracy.

The GRU algorithm, a deep learning technique, addresses the low accuracy problem. It is also employed in speech processing, digital signal processing, language modeling, and human activity recognition. The primary aim of the GRU algorithm is to mitigate the vanishing gradient problem, making it particularly useful in processing image sequences over time.

II. RELATED WORKS

Hammerer NY et al. [14] experimented with research on the Opportunities dataset to identify a wide range of human activities using the Bi-directional LSTM algorithm, with hand signals as input sourced from inertial sensors. The research yielded an F1-score result of about 91%.

Qiu S et al. [15] developed an LSTM algorithm employing regularization techniques on the HARTH dataset, achieving an accuracy of about 94%. In their research, they used two parameters to assess the results: the confusion matrix and the learning curve.

Cruciaini et al. [16] experimented with human activity recognition on the UCI - HAR dataset using Convolutional Neural Networks and achieved an accuracy of about 91.8%. Ordonez et al. [17] proposed a convLSTM technique to recognize and classify human activities. The authors used various initial sensors and accelerometers to measure parameters against the Skoda dataset, classifying around five human activities. The algorithm yielded an F1 score of about 95.8% accuracy.

[1] Assistant Professor, School of Electronics Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India. prasannabharathi.s@vit.ac.in

[2] Assistant Professor, Department of Electronics and Communication Engineering, Saveetha Engineering College, Chennai, Tamil Nadu, India. chamundeeswarig@saveetha.ac.in

[3] Assistant Professor, School of Electronics Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India. upender.p@vit.ac.in

Xia K et al. [18] proposed a CNN Long Short-Term Memory algorithm against the HARTH dataset for recognizing human activities. This algorithm achieved around 95.8% accuracy and required less computational time during the training phase.

Alani et al. [19] proposed CNN, LSTM, and CNN LSTM against the SPHERE dataset, achieving accuracies of 92.89%, 93.55%, and 93.6% respectively. In this research, the CNN, CNN LSTM, and LSTM algorithms were used for imbalanced data for HAR, with limitations in performance.

Alzantot et al. [20] used LSTM techniques to classify human activities and categorize real and synthesized (generated) data. The drawback of this research is low accuracy and longer training times during the training phase due to the presence of many training layers leading to complex architecture. Shakya et al. [21] experimented with CNN and RNN algorithms against the Actitracker dataset, attaining an accuracy of about 92.2%.

III. Proposed Methodology

In this research, a ReLU layer is incorporated along with GRU to enhance classification accuracy in identifying individual activities. The HARTH dataset is utilized for training and testing GRU. The dataset is split into training and testing phases, with 80% of the data used for training and the remainder for testing purposes. Figure 1 illustrates the proposed methodology. In this setup, the GRU algorithm is trained to classify six classes: walking, running, standing, upstairs, downstairs, and sitting.

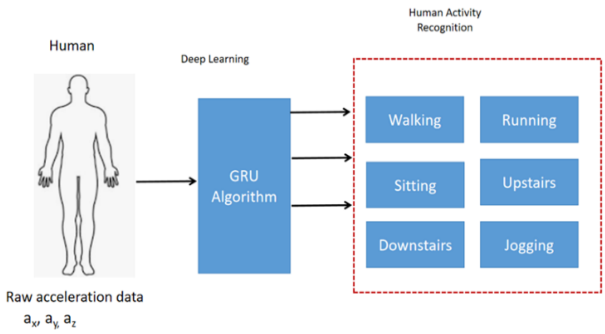

Figure 1: Framework of the proposed system

The Gated Recurrent Unit is one of the deep learning algorithms, a type of recurrent neural network. The vanishing gradient problem in RNN is resolved by the GRU algorithm. It is an algorithm utilized to process sequences of time data. The GRU unit does not have any memory units to get linked with. The architecture of GRU is shown in Figure 2. The GRU consists of a reset gate and an update gate. This setup includes a sigmoid function and based on the output of the sigmoid function, the data is either accepted or rejected by the gate. The information is passed across the cell based on the sigmoid output. When the sigmoid output is 1, data is accepted by the gate, and when the sigmoid output is 0, the data is rejected by the gate. The output of the sigmoid function is denoted as R_t. The input to the Reset gate, h_t, is considered as the present state, and the previous hidden state is given by h_{t-1}.

The sigmoid function is attached to the Update gate, which in turn updates the cell state. The Update gate decides whether to pass or reject the data to the future state via the gate. The sigmoid function ranges between 0 and 1. For all values above zero, the output of the sigmoid function will be 1, whereas for all negative values and zero, the output of the sigmoid is 0. Similarly, we use a tanh activation function along with the sigmoid function to create the new state. The tanh has a value ranging from -1 to 1. The sigmoid function is multiplied with the tanh function to give the output. The output of the present cell C_t is added with the multiplication product, where h_t is the hidden state output of the current cell. Reset and update gate are computed by the following equations, respectively. W_r and W_z represent the reset and update gate's weights. The corresponding weight was calculated during the training phase, respectively.

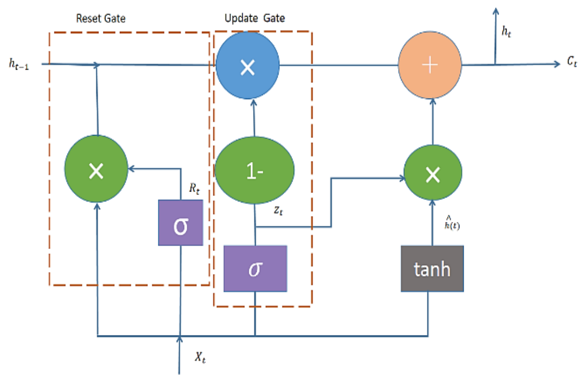

Figure 2: Architecture of GRU cell

Moreover, h_t, the height of the GRU, is given by the following equations.

$$R_t = \sigma(W_r[h_{t-1}, x_t]) \quad - (1)$$

$$Z_t = \sigma(W_z[h_{t-1}, x_t]) \quad - (2)$$

$$h_t = (1 - Z_t)h_{t-1} + Z_t\,\hat{h}_t \quad - (3)$$

The GRU is refined and implemented here to recognize and classify human activities. It is a special kind of RNN, employed to resolve the vanishing gradient problem present in the RNN. Apart from this it has an added advantage, that it requires less computational time in the training phase because of its high speed. It is well suited to process time sequence data and works better than RNN.

Figure 3 represents the overall description of GRU and it consists of an input layer, two GRU layers and a single output layer. The input layer consists of three input features collected from the recordings of the accelerometers positioned in the individuals. a_x, a_y, a_z are considered as the input features, where a_x, a_y and az denotes acceleration in the x, y and z axis respectively.

The input time sequence data is captured by the two GRU layers which is fed by the input layer as a feature input. Each Gated Current Unit layer has 32 hidden layers and a ReLU function along with it. The ReLU function is used for the robustness of GRU. Generally, GRU layers are stacked one above the other to increase the stability and accuracy. The output layer comprises of six neurons to identify six classes. Each neuron has a softmax activation function along with it to identify the classes and also cross entropy function for wrong identification is calculated here in terms of probability function. The learning rate of the algorithm for training the data is set to 0.0030 and each batch length is fixed to 64.

The learning rate decides the training rate of the GRU algorithm. The weights of the GRU algorithm were computed with the help of Adam Optimizer function. The computation of the proper weights for the training process increases the training accuracy, reduces the loss and minimizes the error. Additionally, a regularization method is incorporated to avoid over-fitting in the training phase. Cross entropy loss is calculated to identify the errors of wrong identification of the classes in the overall proposed framework.

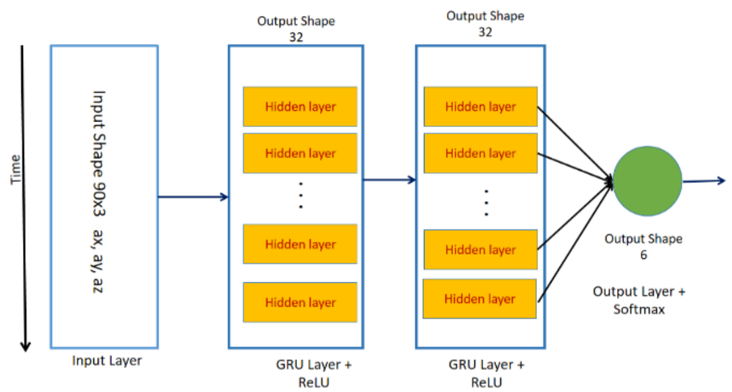

Figure 3: GRU Framework

IV. Dataset Description

The HARTH dataset comprises raw acceleration data collected from 22 individuals who wore two to three axial accelerometers on their bodies to measure the data. The three accelerometers were worn on the lower back and thigh of their bodies, and the corresponding accelerometer signals were measured while they performed any activity. It consists of 12 different classes of activities performed by humans, such as walking, standing, shuffling, climbing steps (up and down), sitting, lying, running, cycling with sitting position active, cycling with standing position active, cycling - sitting idle, and cycling - standing idle. The complete dataset is split into training data (80%) and testing data (20%). The HARTH dataset is purely based on human activities, which are based on the x, y, and z acceleration, time-dependent, and it is a function of gravity (g). The magnitude of the three recorded signals varies from -g to g.

Evaluation metrics are used to measure the model's performance using various statistical parameters from the Confusion Matrix, such as Precision, Sensitivity, Accuracy, and F1-Score. False Negative (FN), False Positive (FP), True Positive (TP), and True Negative (TN) are the measures obtained from the Confusion Matrix. A correctly predicted output from a model is termed as TP, while True Negative (TN) means a model correctly predicts it as negative, also known as correct rejection. FP indicates a wrong class, but the model predicts it as correct (false alarm or overestimation), and FN indicates that the model predicts it as negative, but it is actually positive, which is a type II error (underestimation).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \quad - (4)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad - (5)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad - (6)$$

$$\text{F1 score} = 2 \text{ x } \frac{\text{Precision x Recall}}{\text{Precision + Recall}} \quad - (7)$$

V. Results And Discussion

The algorithm was implemented for the data fed as an input and the obtained confusion matrix for the same is given below.

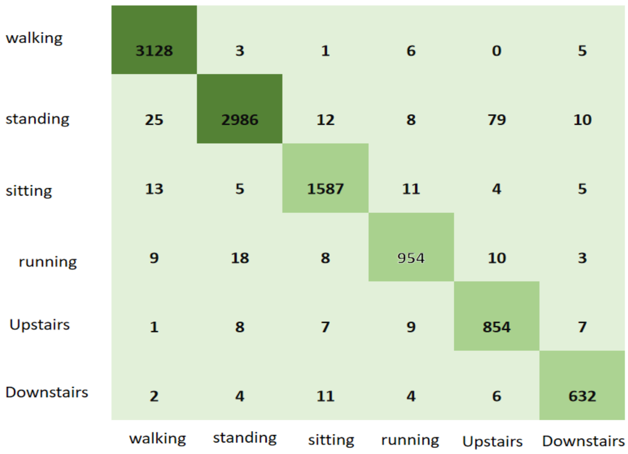

Figure 4: Confusion matrix of the proposed framework against HARTH dataset

The confusion matrix denotes the prediction summary of each class in matrix form. It displays correct and incorrect predictions, providing a better understanding of the different classes present in the dataset.

Figure 4 represents the Confusion matrix of the proposed framework. The proposed system detects 3128, 2986, 1587, 954, 854 and 632 for the 6 activities performed by human such as walking, standing, sitting, running, Upstairs and downstairs correspondingly. The Precision, F1 – Score, recall for all the 6 classes have been tabulated below.

Table 1 Classification Summary for the proposed GRU framework

Class	Precision	Recall	F1 - Score
Walking	0.9531	0.9512	0.9527
Standing	0.9234	0.9287	0.9210
Sitting	0.9711	0.9797	0.9738
Running	0.9034	0.9089	0.9013
Upstairs	0.9514	0.9587	0.9598
Downstairs	0.9423	0.9419	0.9479

Table 2 Comparison of accuracy of the previous works with the proposed framework

Reference	Algorithm	Dataset	Accuracy (%)
Cruciani Et al.	CNN	UCI -HAR	91.98
Hammerla Et al	Bidirectional	Opportunity	92.7
LSTM			
Pienaar Et al	LSTM	WISDM	94
Current Work	GRU	HARTH	94.98

Table 1 indicates the values obtained for Precision, Recall and F1-Score for all the six classes.
Table 2 shows the proposed framework implemented across HARTH dataset and the comparative analysis of the results obtained using other datasets.

VI. CONCLUSION

The GRU algorithm is implemented across the HARTH dataset to detect human activities and its classification accuracy was measured about 94.8%. Other evaluation parameters like Precision, Sensitivity (Recall) and F1 Score had also found and reported that GRU algorithm outperforms the previous works such as CNN and other machine learning algorithms. In future, hyper parameter tuning can be implemented to increase classification accuracy furthermore. In addition, it can be implemented with other deep learning algorithms for better results.

REFERENCES

[1] Mohsen S, Elkaseer A, Scholz SG, "Human activity recognition using k-nearest neighbor machine learning algorithm," Proceedings of the 8th International Conference on Sustainable Design and Manufacturing (KES-SDM), vol. 262, pp. 304–313, 2021.

[2] Demrozi F, Bacchin R, Tamburin S, Cristani M, Pravadelli G "Towards a wearable system for predicting freezing of gait in people affected by Parkinson's disease", IEEE Journal of Biomedical Health Informatics, vol.24, pp. 2444–2451, 2019.

[3] Hsu Y, Chang H, Chiu Y, "Wearable sport activity classification based on deep convolutional neural network," IEEE Access, vol. 7, pp. 170199–170212, 2019.

[4] Lawal IA, Bano S, "Deep human activity recognition with localization of wearable sensors," IEEE Access, vol. 8, pp. 155060–155070, 2020.

[5] Malaisé A, Maurice P, Colas F, Charpillet F, Ivaldi S, "Activity recognition with multiple wearable sensors for industrial applications," Proceedings of the 1st International Conference on Advances in Computer Human Interactions, pp. 1–7, 2019

[6] Tanberk S, Kilimci ZH, Tükel D, Uysal M, Akyokuş S, "A hybrid deep model using deep learning and dense optical flow approaches for human activity recognition," IEEE Access vol. 8, pp. 19799–19809, 2020.

[7] Wang Y, Cang S, Yu H, "A survey on wearable sensor modality centered human activity recognition in healthcare," Expert System Applications, vol. 137, pp. 167–190, 2019.

[8] Lattanzi E, Freschi V, "Evaluation of human standing balance using wearable inertial sensors: a machine learning approach," Engineering Application Artificial Intelligence, vol. 94, pp. 103812, 2020.

[9] Ronao C A, Cho S B, "Human activity recognition with smartphone sensors using deep learning neural networks," Expert System Applications, vol. 59, pp.235–244, 2016.

[10] Shoaib M, Bosch S, Incel OD, Scholten H, Havinga PJM, "A survey of online activity recognition using mobile phones," Sensors vol. 15, pp. 2059–2085, 2015.

[11] Wang J, Chen Y, Hao S, Peng X, Hu L, "Deep learning for sensor-based activity recognition: A Survey," Pattern Recognition Letters, vol. 119, pp.3–11, 2018.

[12] Hong J, Ramos J, Dey AK, "Toward personalized activity recognition systems with a semi population approach," IEEE Transactions Human Machine Systems, vol. 46, pp. 101–112, 2016.

[13] Igual R, Medrano C, Plaza I, “A comparison of public datasets for acceleration-based fall detection,” Medical Engineering Physics, vol. 37, pp. 870–878, 2015.

[14] Hammerla NY, Halloran S, Plötz T, “Deep, convolutional, and recurrent models for human activity recognitionaion using wearables,” Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, pp. 1533–1540, 2016.

[15] Qiu S, Zhao H, Jiang N, Wang Z, Liu L, An Y, Zhao H, Miao X, Liu R, Fortino G, “Multi-sensor in formation fusion based on machine learning for real applications in human activity recognition: state-of-the-art and research challenges,”. Information Fussion, vol. 80, pp.241–265, 2022

[16] Cruciani F, Vafeiadis A, Chris Nugent, Ian Cleland, Paul McCullagh, Konstantinos Votis, Dimitrios Giakoumis, Dimitrios Tzovaras, Liming Chen, Raouf Hamzaoui, “Feature learning for human activity recognition using convolutional neural networks,” CCF Transactions on Pervasive Computing and Interaction, vol. 2, pp. 18–32, 2020.

[17] Ordóñez FJ, Roggen D, “Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition,” Sensors vol. 16, 2016.

[18] Xia K, Huang J, Wang H, “LSTM-CNN architecture for human activity recognition. IEEE Access, vol. 8, pp. 56855–56866, 2020.

[19] Alani AA, Gosma G, Taherkhani A, “Classifying imbalanced multi-modal sensor data for human activity recognition in a smart home using deep learning,” International Joint Conference on Neural Networks (IJCNN) pp. 1–8, 2020.

[20] Alzantot M, Chakraborty S, Srivastava M, “SenseGen: A deep learning architecture for synthetic sensor data generation,” IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), pp. 188–193, 2017.

[21] Shakya SR, Zhang C, Zhou Z, “Comparative study of machine learning and deep learning architecture for human activity recognition using accelerometer data,” International Journal of Machine Learning and Computing, vol. 24, pp. 292–299, 2018.